

REPORT DOCUMENTATION PAGE			Form Approved OMB NO. 0704-0188	
<small>Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comment regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.</small>				
1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE 20 October 1998		3. REPORT TYPE AND DATES COVERED Final Progress - Sept. 1, 95-Aug. 31, 98
4. TITLE AND SUBTITLE Multiresolution, Multi-Scale Target Identification and Tracking using the Anisotropic Diffusion Pyramid			5. FUNDING NUMBERS DAAH04-95-1-0255	
6. AUTHOR(S) Scott T. Acton, Associate Professor				
7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(ES) Oklahoma State University, Stillwater, Oklahoma 74078			8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Research Office P.O. Box 12211 Research Triangle Park, NC 27709-2211			10. SPONSORING / MONITORING AGENCY REPORT NUMBER  ARO 3447.20-EL-YIP	
11. SUPPLEMENTARY NOTES The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.				
12a. DISTRIBUTION / AVAILABILITY STATEMENT  Approved for public release; distribution unlimited.			12 b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words) A three year ARO Young Investigator project has resulted in a novel target identification and tracking system based on multiresolution, multi-scale image processing methods. Through this research effort, two nonlinear image processing methods have been developed and utilized in target tracking: the anisotropic diffusion pyramid and the morphological pyramid. Coarse-to-fine target searches are implemented within the image pyramids, providing a 100X improvement in computational expense over standard correlation-based approaches. Several improvements in tracking robustness have been achieved, including a 10X improvement in target localization over standard linear processing methods. A noise-resilient morphological diffusion method has been designed, and diffusion techniques that converge rapidly to locally monotonic signals have been developed, yielding a significant advance in nonlinear diffusion-based image processing. Further improvements to the diffusion technique have included the incorporation of multigrid methods and the generalization to multispectral imagery for multisensor tracking applications. These advances in theory have been validated with a new tracking simulator, in which several visible-wavelength and infrared image sequences have been used for testing. During the last three years, the ARO-funded project has supported 10 graduate students at Oklahoma State University and has led to over 20 refereed publications.				
14. SUBJECT TERMS Target tracking, automated target recognition, image processing			15. NUMBER OF PAGES 46	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT UNCLASSIFIED	18. SECURITY CLASSIFICATION OF THIS PAGE UNCLASSIFIED	19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED	20. LIMITATION OF ABSTRACT UL	

**MULTIRESOLUTION, MULTI-SCALE TARGET IDENTIFICATION AND  
TRACKING USING THE ANISOTROPIC DIFFUSION PYRAMID**

**FINAL PROGRESS REPORT**

**SCOTT T. ACTON**

**OCTOBER 20, 1998**

**U.S. ARMY RESEARCH OFFICE**

**DAAH04-95-1-0255**

**OKLAHOMA STATE UNIVERSITY**

**APPROVED FOR PUBLIC RELEASE;**

**DISTRIBUTION UNLIMITED**

19981230 034

**THE VIEWS, OPINIONS, AND/OR FINDINGS CONTAINED IN THIS REPORT ARE  
THOSE OF THE AUTHOR AND SHOULD NOT BE CONSTRUED AS AN  
OFFICIAL DEPARTMENT OF THE ARMY POSITION, POLICY, OR DECISION,  
UNLESS SO DESIGNATED BY OTHER DOCUMENTATION.**

## TABLE OF CONTENTS

A. STATEMENT OF THE PROBLEM STUDIED .....	1
B. SUMMARY OF THE MOST IMPORTANT RESULTS .....	3
Scale Space.....	3
Multi-scale Image Pyramids.....	5
Anisotropic Diffusion.....	6
Morphological Anisotropic Diffusion.....	11
Anisotropic Diffusion Pyramids.....	27
Multi-scale Tracking .....	28
Experimental Results .....	32
C. PUBLICATIONS .....	39
D. PARTICIPATING PERSONNEL AND DEGREES AWARDED.....	41
E. REPORT OF INVENTIONS.....	41
BIBLIOGRAPHY .....	42

## LIST OF FIGURES

FIGURE 1 .....	3
FIGURE 2 .....	4
FIGURE 3 .....	5
FIGURE 4 .....	7
FIGURE 5 .....	8
FIGURE 6 .....	10
FIGURE 7 .....	12
FIGURE 8 .....	13
FIGURE 9 .....	14
FIGURE 10 .....	16
FIGURE 11 .....	16
FIGURE 12 .....	17
FIGURE 13 .....	17
FIGURE 14 .....	18
FIGURE 15 .....	19
FIGURE 16 .....	21
FIGURE 17 .....	24
FIGURE 18 .....	25
FIGURE 19 .....	26
FIGURE 20 .....	29
FIGURE 21 .....	31
FIGURE 22 .....	32
FIGURE 23 .....	33
FIGURE 24 .....	34
FIGURE 25 .....	34
FIGURE 26 .....	35
FIGURE 27 .....	36
FIGURE 28 .....	36
FIGURE 29 .....	37
FIGURE 30 .....	37
FIGURE 31 .....	38
FIGURE 32 .....	38

## **A. STATEMENT OF THE PROBLEM STUDIED**

The problem studied in this ARO project is automated target identification and tracking using efficient multiresolution image processing techniques. Advanced forward-looking infrared (FLIR) systems are capable of producing very high-resolution imagery at large frame rates and generating an enormous amount of raw image data. Infrared search and track (IRST) procedures must process this information reliably in a variety of battlefield environments and imaging situations. Corruptive image noise introduced by dust, smoke or other clutter can degrade the performance of an IRST system. Vibration effects can introduce image blurring, an increase in the minimum detection temperature (MDT), amplification of high frequency noise, reduced display comprehensibility, and reduced signal-to-noise ratios (SNR) (Miller, 1993). Furthermore, bugs, salt, leaves, fuel and moisture may obstruct the FLIR protective window.

Imaging under battlefield conditions dramatically increases the difficulty of the IRST task. In addition to environmental conditions, signal processing algorithms are further constrained to operate with minimal computational resources in order to reduce the overall weight and size of the system package. Given modern sensors that are capable of producing raw data at rates of over one million pixels per second, traditional fixed resolution approaches introduce exceptional computational burden. At this data rate, the fixed resolution (correlation based) approach to automatic target recognition requires over a billion operations per second and precludes real time processing (Nasr, 1989; Molley, 1989).

The majority of detail processed by fixed resolution imaging systems is irrelevant for the IRST task. Small scale features and channel noise increase the computational requirements of the system but introduce little information for improved target detection. Additionally, the discrimination and extraction of region boundaries for potential targets are complicated by the higher resolution imagery. For improved IRST performance and decreased computational requirements, a hierarchical approach should be employed. These architectures mimic biological vision systems by initially searching coarse scale scene representations. These coarse scale results may then be exploited to efficiently process finer resolution data. For example, humans initially identify peripheral objects as potential regions of interest, acquiring higher-resolution scene information by focusing on the region and then deciding if the perceived object is present.

Biological search procedures are facilitated by the nonlinear distribution of visual sensors within a biological vision system. However, the majority of FLIR sensors do not utilize nonlinear sensor distributions. Instead, a foveating IRST system must replicate the advantages of a nonlinear scene description with uniformly sampled data. This concept is encapsulated in continuous *scale space theory* (Witkin, 1983), where it is proposed that an infinite number of coarse scene representations may be created by filtering the original imagery with a linear, *scale generating*, filter. The subsequent data structure can then be queried in a manor analogous to the biological coarse-to-fine search, as objects are initially identified in coarse scale scene representations that are absent of small scale clutter, fine features, texture and noise. These initial coarse scale results are then used to guide and refine higher resolution inspection, a process that terminates with the identification of features in the original imagery.

The application of scale space theory to a practical IRST system is problematic. Construction of scale space requires a large number (theoretically infinite) of scale representations to follow features from coarse scale information to finer scene depictions. Additionally, inspecting coarse scale imagery is computationally equivalent to fixed resolution searches, as the original and coarse scale descriptions are stored at equivalent sample densities. In application, these characteristics increase storage and computational requirements and result in added system weight and power consumption. With limited resources, direct use of scale space data structures in a battlefield environment is presently unfeasible. Employing the robust properties of scale space in a resource critical IRST system necessitates the quantization (via sampling) of scale space.

This report provides details on the engineering solution achieved during the course of the ARO-sponsored research. Two multiresolution structures for image processing, the anisotropic diffusion pyramid and the morphological pyramid, were developed and utilized in target tracking. Several important advances were achieved in diffusion-based processing. The anisotropic diffusion methods were extended to multigrid and multispectral implementations, allowing multisensor tracking. The parameter selection processes were automated and the diffusion methods were made more robust through morphological filtering. The tracking simulation results show substantive improvements in both solution time and solution quality.

## **B. SUMMARY OF THE MOST IMPORTANT RESULTS**

### ***Scale Space***

Scale space filtering was initially developed to manage the relationship between edge information over varying resolution. Since many signal characteristics, most notably derivatives, are calculated over a region where the region size influences the descriptive measurement, Witkin introduced scale space as a collection of signal representations, derived from the original image and generated by a scale space filter (Witkin, 1983). Scale space does not attempt to define an optimal scale for feature identification but provides a method for establishing correspondence between edges found in heavily filtered signal representations and their location in the original signal.

Construction of scale space is straightforward, and traditionally begins by filtering the original signal with an FIR filter of varying width. Hyper-planes within scale space contain a single filtered representation of the signal, while filtering the signal with a continuum of filter widths produces scale space. For a two-dimensional image, scale space may be visualized as a three-dimensional cube, containing an infinite number of signal descriptions stacked upon

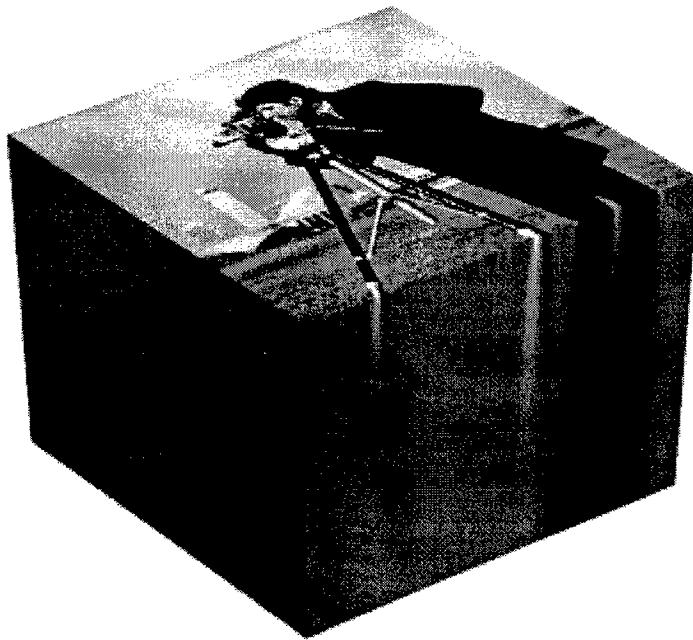


Figure 1. Scale space for the cameraman image. The original image is located at the top of the cube and lower levels are occupied by coarser representations of the scene. For this example, the coarse scale images are constructed by smoothing the original image with a Gaussian filter.

one another. These representations are ordered by their respective filter scale parameter. From these filtered images, features may be identified using traditional detection methods. An example scale space is shown in Figure 1.

Plotting the locations of detected features versus the continuous scale parameter is defined as the *scale space image* and an example is displayed in Figure 2. Within these structures, objects may be recognized at coarse resolution representations and traced to their origin. This is referred to as a *coarse to fine search* and encompasses the power of scale space theory, allowing the initial identification of significant features to occur in the absence of spurious derivative results. The exact location of these edge points in the original image may then be obtained by traversing scale space towards finer resolution, resulting in a robust method for fusing multi-scale information and a procedure well suited to the edge detection problem.

Specification of a viable scale space filter requires fulfilling a specific smoothing criteria: if a feature is tracked across increasing scale, it should disappear. Conversely, a new feature should never appear while scale increases, as coarse resolution representations would no longer correspond to the original signal. Guaranteeing the presence of coarse scale objects in finer scene representations is expressed as *spatial causality*, maintaining a cause and effect

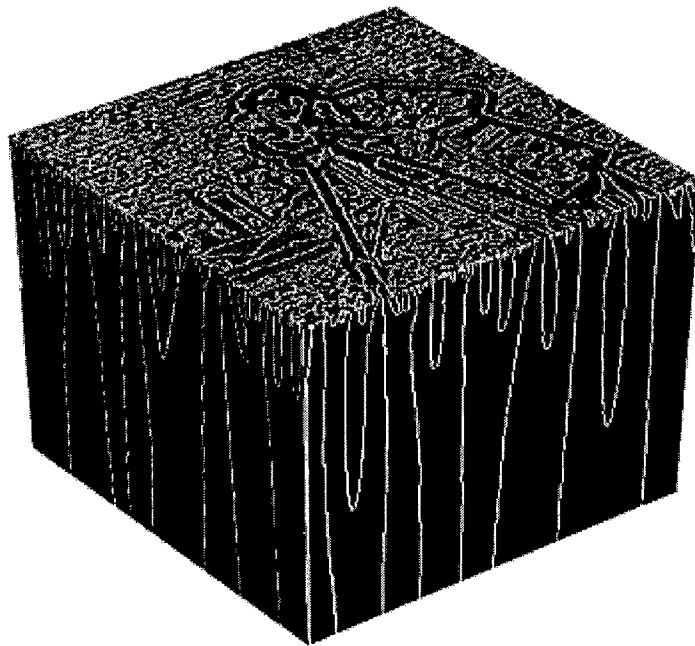


Figure 2. Scale space and the scale space image. Scale space images are displayed on the sides of the cube and show how smaller features disappear rapidly as scale is increased. Connectivity between levels is defined as spatial causality, as all coarse scale features correspond to features in the higher resolution representations.

relationship for features, and is a necessary condition for application of the multi-scale coarse to fine search method.

The spatial causality criterion allows the specification of an optimal filter for scale space generation. Witkin initially restricted the scale generating filter to be symmetric, strictly decreasing about the mean, and linear. As a result of this definition, it has been shown that the Gaussian kernel is the only filter capable of satisfying these constraints in one-dimension while maintaining spatial causality (Babaud, 1986). The uniqueness of the Gaussian kernel for scale space construction has since been extended to higher dimensions (You *et al.*, 1996), discrete signals (Lindeberg, 1990), and the larger class of unsmooth kernels (Wu, 1990). The Gaussian filter also has the unique property of minimizing the uncertainty principle (Marr, 1980).

Even with the discovery of a unique scale generating filter, application of scale space theory to practical problems is limited. Requiring an infinite (or near infinite) number of scale representations necessitates large storage requirements, and performing feature detection tasks on each resolution level is computationally expensive. Efficient execution of a coarse to fine search demands quantization of the scale parameter, and is formalized by the image pyramid.

### **Multi-scale Image Pyramids**

Image pyramids are a discrete representation of scale space. By requiring the calculation of fewer scene representations, they reduce the computational requirements of scale space construction and the coarse-to-fine search. Image pyramids also introduce additional processing speed-up by coupling their choice of scale retention to the sampling properties of a scale generating operator. Allowing the decimation of coarse resolution representations results in decreased storage requirements, faster scale space construction, and a logarithmic improvement in coarse-to-fine matches. Theoretically, an image pyramid will provide a very efficient *and* robust solution to the IRST problem.

Construction of an image pyramid begins by filtering the original signal. This coarser resolution representation, now satisfying some sampling criterion, is then decimated. The sampling process traditionally consists of discarding all pixels belonging to the even rows and columns of the image. Subsequent pyramid levels are created by iteratively filtering and

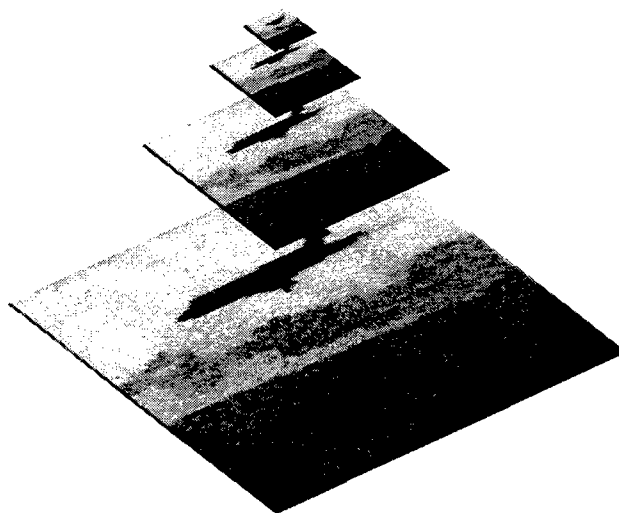


Figure 3. An image pyramid, constructed by filtering the original image and subsampling. The original image is located at the bottom of the pyramid and coarser scale representations occupy successively higher pyramid levels. For this example, the coarse scale images are constructed by smoothing the original image with a Gaussian filter.

subsampling the previous resolution representation, and the end product is a set of image descriptions, each of smaller size than the original.

For example, Gaussian pyramids are constructed by using a Gaussian kernel as the scale generating filter and applying Shannon's sampling theorem for the decimation operation. Mathematically, the construction of pyramid level  $L$  of image  $\mathbf{I}$  can be described as

$$\mathbf{I}_L = [\mathbf{G}_\sigma * \mathbf{I}_{L-1}] \downarrow_S \quad (1)$$

where  $\mathbf{G}_\sigma$  is a Gaussian scale generating filter of standard deviation  $\sigma$ ,  $\downarrow_S$  denotes subsampling by a factor of  $S$  within each row and column, and  $\mathbf{I}_0$  is the original image. To remove frequencies below the Nyquist rate, the discrete Gaussian filter is defined to have kernel width  $2S/\pi$  (Burt, 1981). Sorting these images according to scale is shown graphically in Figure 3 and presents the visual appearance of a pyramid structure, hence the name image pyramid.

Generation of an image pyramid provides significant performance enhancement for the coarse-to-fine search. Initialized on the coarse resolutions in the pyramid, the search procedure needs to only fully search a subsampled image representation. The result of this first search is then used to guide and refine progressively higher resolution inspections, restricting these subsequent examinations to small regions within the next scale description. Ignoring the effort for pyramid construction, the increase in computational cost of searching the full resolution image, as compared to the multi-scale search, has been shown to be  $S^{2L}$ , where  $S$  is the sample factor and  $L$  is the coarsest level of the search (Wong, 1978). The performance improvement between an image pyramid and traditional scale space would be the same, actually amplified by the decreased construction costs of an image pyramid.

Searching an image pyramid requires *causality* to hold between scale representations. Features that exist at coarse scale must correspond to higher resolution objects. Among the linear kernels, the Gaussian filter is the only operator that maintains spatial causality (Babaud, 1986). Unfortunately, the Gaussian kernel possesses several undesirable characteristics for quantizing the theory of scale space within an IRST system. Specifically, as the scale parameter (the width of the Gaussian) of the filter is increased, regions tend to merge and edges move due to the lowpass response of the filter. In continuous scale space, movement between resolution representations presents no obstacle to the coarse-to-fine search and is accommodated by allowing minimal scale change between neighboring levels. Image pyramids contain a limited number of scale representations, reducing the number of scales available to a multi-scale IRST procedure. With fewer scale depictions, large feature movement between pyramid levels is possible. This source of error dramatically decreases system robustness and performance. In order to attain the theoretical promise of multi-scale search and track, nonlinear scale generating operators must be considered. An ideal operator would describe a scale space with minimum feature movement, as important objects maintain the same spatial location independent of scale. One scale generating filter, *anisotropic diffusion*, can be designed specifically for this task.

### **Anisotropic Diffusion**

Anisotropic diffusion is a departure from traditional linear filtering. Linear filters are widely used in signal processing and are theoretically well developed. The anisotropic diffusion equation modifies the behavior of one member of the linear filter class, the

Gaussian kernel. In the traditional scale space representation, a Gaussian pyramid is usually constructed by convolving the original image with a suitable Gaussian kernel and subsampling. This multi-scale structure can also be implemented with the use of the heat equation (Koenderink, 1984). For an image defined on a discrete grid, this process is approximated by the following partial differential equation (PDE):

$$\frac{\partial \mathbf{I}}{\partial t} = \nabla^2 \mathbf{I}, \quad (2)$$

where  $\mathbf{I}$  is the input signal,  $\nabla^2$  is the discrete Laplacian, and  $t$  is the solution time or scale parameter.

Anisotropic diffusion modifies the smoothing properties of the Gaussian kernel by adaptively smoothing within regions while inhibiting intra-region interaction. Solution of the heat equation is completely defined by its Green's function, the Gaussian kernel, with the width of the resulting kernel proportional to solution time (Widder, 1975). In creating a scale generating process with the capability of maintaining region integrity, the heat equation may be modified to incorporate a spatially varying damping coefficient. The PDE becomes

$$\frac{\partial \mathbf{I}}{\partial t} = \text{div}(\mathbf{c} \cdot \nabla \mathbf{I}), \quad (3)$$

where  $\mathbf{I}$  is the input signal,  $\text{div}$  is the divergence operator,  $\nabla$  is the discrete gradient, and  $\mathbf{c}$  is the adaptive diffusion coefficient. For a two dimensional image, one possible realization is

$$\mathbf{I}_{t+\Delta t} = \mathbf{I}_t + \Delta t \cdot (c_N \nabla_N \mathbf{I} + c_S \nabla_S \mathbf{I} + c_E \nabla_E \mathbf{I} + c_W \nabla_W \mathbf{I}) \quad (4)$$

where  $\Delta t$  is the solution time step,  $\nabla_N$ ,  $\nabla_S$ ,  $\nabla_E$ , and  $\nabla_W$  are the gradients (simple differences) in the north, south, east, and west directions, respectively, and  $c_N$ ,  $c_S$ ,  $c_E$ , and  $c_W$  are the diffusion coefficients in the north, south, east, and west directions, respectively (Perona, 1990). These coefficients are traditionally bounded to the set  $[0,1]$  and decrease with

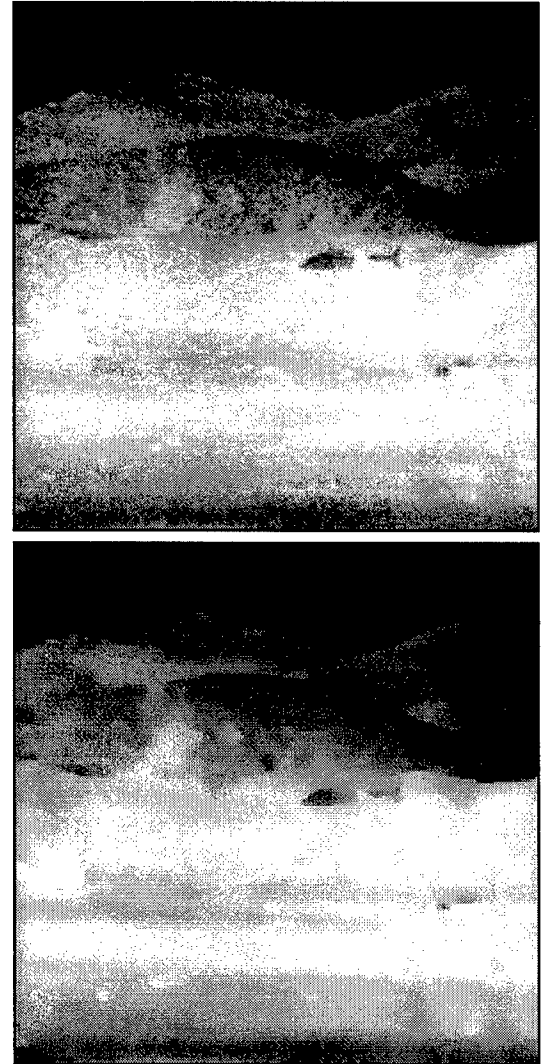


Figure 4. A visual example of filtering with the anisotropic diffusion equation. The original image is located at the top, and its filtered result is located below. Notice how the diffusion process smooths within the boundaries of an object but preserves edge locations.

increasing gradient, effectively inhibiting smoothing in regions of possible edge location. An example of filtering an image with the anisotropic diffusion equation is presented in Figure 4.

Construction of the anisotropic diffusion coefficient defines the performance of diffusion. With the initial goal of preserving regions of possible transition, varying the coefficient relative to local gradient is well motivated. In the first use of anisotropic diffusion as a filtering process, two possible realizations of the diffusion coefficient were suggested (Perona, 1990). The first is of Gaussian shape and expressed as

$$c(\nabla I) = e^{-\left(\frac{\|\nabla I\|}{k}\right)^2}, \quad (5)$$

while the second suggestion is

$$c(\nabla I) = \frac{1}{1 + \frac{\|\nabla I\|}{k}}. \quad (6)$$

In both diffusion coefficients, a gradient threshold,  $k$ , is introduced and its selection quantifies the minimum gradient magnitude which should be preserved by the smoothing mechanism.

In creating an image pyramid using anisotropic diffusion, one would successively diffuse and then subsample the original image. Unfortunately, anisotropic diffusion does not satisfy requirements for image pyramid construction, as nonlinearly filtered signals do not satisfy traditional sampling theorems. Without the assistance of an image pyramid, creating coarse scene information computationally expensive and searching these scale representations provides no performance increase, as the resolution of the original and coarse scene information are identical. More problematic for the application of anisotropic diffusion is that the diffusive process does not guarantee the removal of *any* information. For example, consider the one dimensional pulse train shown in Figure 5, where the pulse heights are identical and defined to be greater than the gradient threshold,  $k$ , of the diffusion coefficient. Anisotropic diffusion is designed to preserve regions of high gradient, and the traditional definitions for damping coefficients result in coefficient values near zero at the

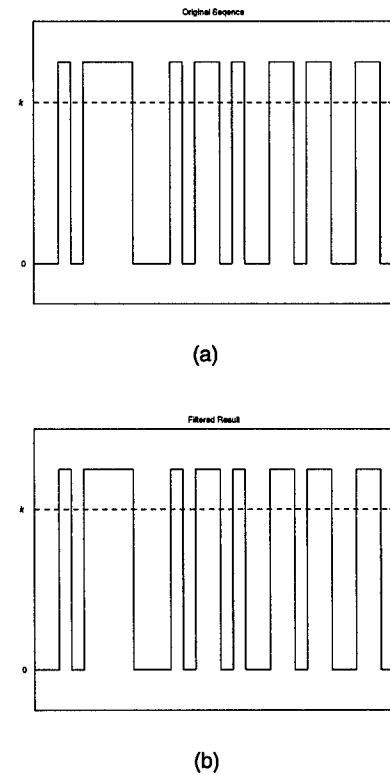


Figure 5. An input sequence that will not be smoothed by the diffusion process: (a) the original sequence and (b) it's filtered result. The gradient threshold of the diffusion system is represented by  $k$ , and all gradients larger than this threshold are defined to be preserved.

pulse edges. Since the diffusion update depends on a weighted sum of the product of local coefficient and gradient magnitudes, smoothing will not occur. The filtered result will be identical to the original, and changes in solution time, or scale, will not effect information removal. At best, scale spaces constructed with traditional anisotropic diffusion present a robust solution to the IRST problem, but at *increased* computational cost.

IRST filtering problems are motivated by the need to remove noise and other spurious detail. Instead, traditional anisotropic diffusion preserves all regions with high contrast. This characteristic introduces difficulties for the application of anisotropic diffusion to any filtering problem, and researchers have spent time addressing it. The result is a modification to the diffusion coefficient, creating a diffusion equation that is *spatially aware*. In these modified diffusion expressions, the goals of the nonlinear smoothing process are expanded, seeking to preserve features of high gradient *and* to remove regions of small spatial scale.

### Spatially Aware Anisotropic Diffusion

Increasing the scope of the diffusion coefficient calculation enlarges the scale of a diffusion equation. By incorporating greater spatial knowledge of the signal into the decision to diffuse, the filtering process is allowed to smooth small regions regardless of local contrast. A method providing the anisotropic diffusion equation with a direct specification of scale was first proposed by Catté *et al.* (Catté, 1992), who suggest utilizing a Gaussian kernel to spatially expand the coefficient computation. Using the original coefficient expression in (5) as an example, a possible spatially aware diffusion coefficient is specified as

$$c(\nabla I) = e^{-\left(\frac{\|G_\sigma * \nabla I\|}{k}\right)^2}, \quad (7)$$

where  $G_\sigma$  is a Gaussian kernel with standard deviation  $\sigma$ .

Other linear filters have been proposed to replace the Gaussian kernel in (7) (Torkamani-Azar, 1996), and proper selection of a spatially aware anisotropic diffusion coefficient is usually motivated by underlying assumptions of the noise distribution within the original signal. While the use of a linear filter within the diffusion coefficient may be viewed as “against the spirit of anisotropic diffusion” (You *et al.*, 1996), initial results display their ability to remove small regions of high contrast. Figure 6 presents a visual example of the smoothing performance of a scale aware anisotropic diffusion process, showing that filters using the spatially enlarged diffusion coefficients are capable of smoothing, and eventually removing, noise.

A major obstacle in constructing image pyramids with spatially aware anisotropic diffusion implementations is that these new smoothing operators embody conflicting definitions of scale. Traditional diffusion equations contain a single scale parameter, corresponding to solution time. Spatially aware anisotropic diffusion operations incorporate a second scale parameter, describing the region over which a diffusion coefficient is calculated. This second parameter also effects the gradient magnitudes maintained by the nonlinear filter. The result is increased feature movement between scale representations and additional inefficiencies in a coarse-to-fine search. The difficulties are overcome in the next section, where a novel morphological diffusion coefficient is discussed. This new diffusion operator is capable of simultaneously smoothing a signal while maintaining edge locations.

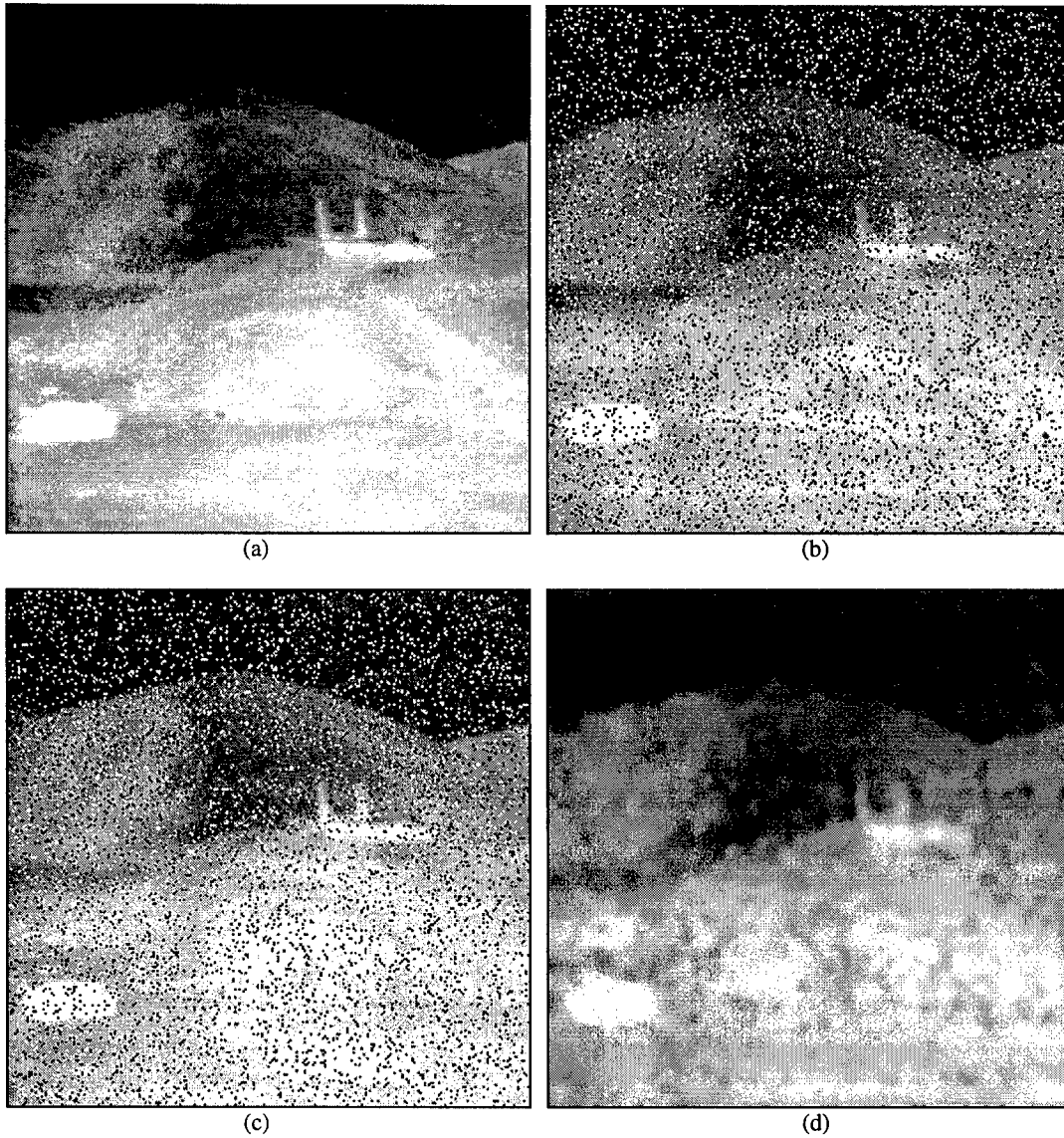


Figure 6. A visual example of filtering with the spatially aware anisotropic diffusion equation: (a) original imagery; (b) imagery corrupted with 15% salt and pepper noise; (c) corrupted imagery after traditional anisotropic diffusion; (d) corrupted imagery after spatially aware anisotropic diffusion. Traditional anisotropic diffusion is unable to remove environmental clutter, small features, texture or noise. Spatially aware anisotropic diffusion is capable of smoothing small regions of high contrast, as evident in the removal of the impulse noise. Spatial smoothing does reintroduce interactions across region boundaries, and this will be discussed in the next section.

### **Morphological Anisotropic Diffusion**

Anisotropic diffusion was originally designed to generate scale spaces with minimal feature movement. Construction of scale spaces with this property alleviates problems in applying coarse-to-fine search methods to image pyramids. Image pyramids, however, require a scale generating process that guarantees the removal of information and satisfies necessary sampling conditions. While traditional anisotropic diffusion expressions are incapable of generating these filtered results, scale aware extensions of the anisotropic diffusion equation attempt to smooth regions of low contrast and small spatial size. A smoothing operator possessing these characteristics would generate a filtered result suitable for sampling and an image pyramid suitable for multi-scale IRST methods.

Initial scale aware realizations of the anisotropic diffusion equation incorporate linear filters into the diffusion coefficient and effectively increase the scope of the coefficient calculation, allowing the smoothing of small regions. However, these expressions are unable to smooth regions of small gradient while removing small scale features. Scale aware diffusion requires regions to be identified without removing their high frequency content. This criterion suggests the use of nonlinear filters in increasing the scope of the coefficient calculation.

Morphological operators are able to produce image representations of increasing scale without eradicating edges. Approaching image processing from the vantage point of human perception, morphological operators simplify image data, preserve essential shape characteristics and eliminate irrelevancies (Toet, 1989; Haralick, 1987). The use of mathematical morphology in digital signal processing is defined with two fundamental operators - *erosion* and *dilation*. Implementation of the erosion and dilation filters is similar to a median filter and is accomplished with nonlinear minimum and maximum operations. An erosion removes regions of high intensity and is expressed as

$$\mathbf{I} \ominus \mathbf{M} = \min_{y \in \mathbf{M}} \{\mathbf{I}(x - y)\}, \quad (8)$$

where  $\mathbf{M}$  is the structuring element and  $\ominus$  is the erosion operator. The mathematical dual of erosion is dilation and removes regions of low contrast by computing the maximum value within a region. Dilation is expressed as

$$\mathbf{I} \oplus \mathbf{M} = \max_{y \in \mathbf{M}} \{\mathbf{I}(x - y)\}, \quad (9)$$

where  $\oplus$  is the dilation operator. In both fundamental filters, the realization of the structuring element defines the shape and scale of the morphological filter and conceptually denotes the signal region from which the minimum (or maximum) value is drawn.

Developing an equivalent gradient representation of the morphological operators furthers understanding of these nonlinear filters. Analysis begins with the step function, which has been shown to be an eigenfunction of the morphological system (Maragos, 1995). The effect of the morphological filters on the step function is shown in Figure 7. Notice that the resulting functions are not smoothed representations of the signal and, instead, are simply shifted by half the width of the structuring element. Applying an erosion to the edge function translates the signal to the right. Alternatively, filtering with a dilation shifts it to the left. The sequential application of these filters results in an infinite number of paths for the step

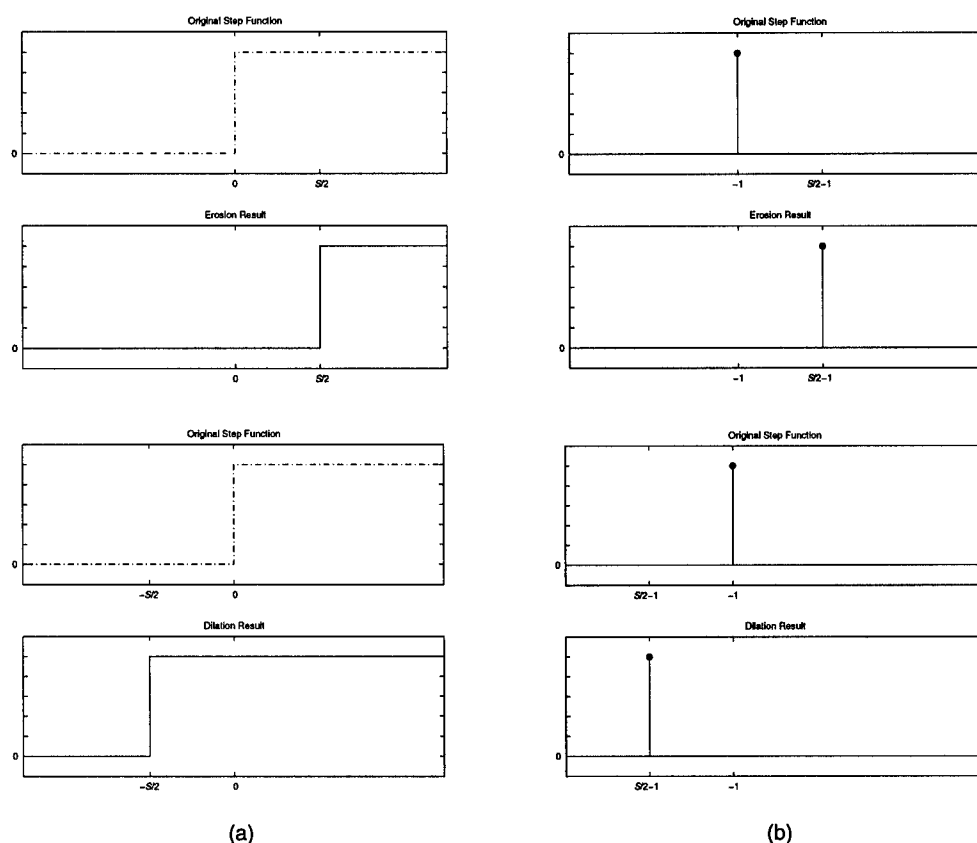


Figure 7. Effect of the fundamental morphological filters on a step function. The equivalent gradient representations are shown in (b).

function gradient to travel, but never modifies the gradient magnitude. Figure 7b displays the right-hand derivatives of the function and its filtered results, again showing the movement of the edge.

A complete gradient understanding of the morphological operators continues with investigation of the negative of the original step function. The signal and its filtered results are displayed in Figure 8. Again, the morphological filters produce translation of the step edge and do not effect the signal amplitude. An important observation is that while the erosion translates the positive edge of Figure 7 to the right, it translates the negative edge in Figure 8 to the left. The dilation produces a similar result, transporting positive and negative edges in opposite directions. This property of sign dependent translation defines the performance of the morphological filters.

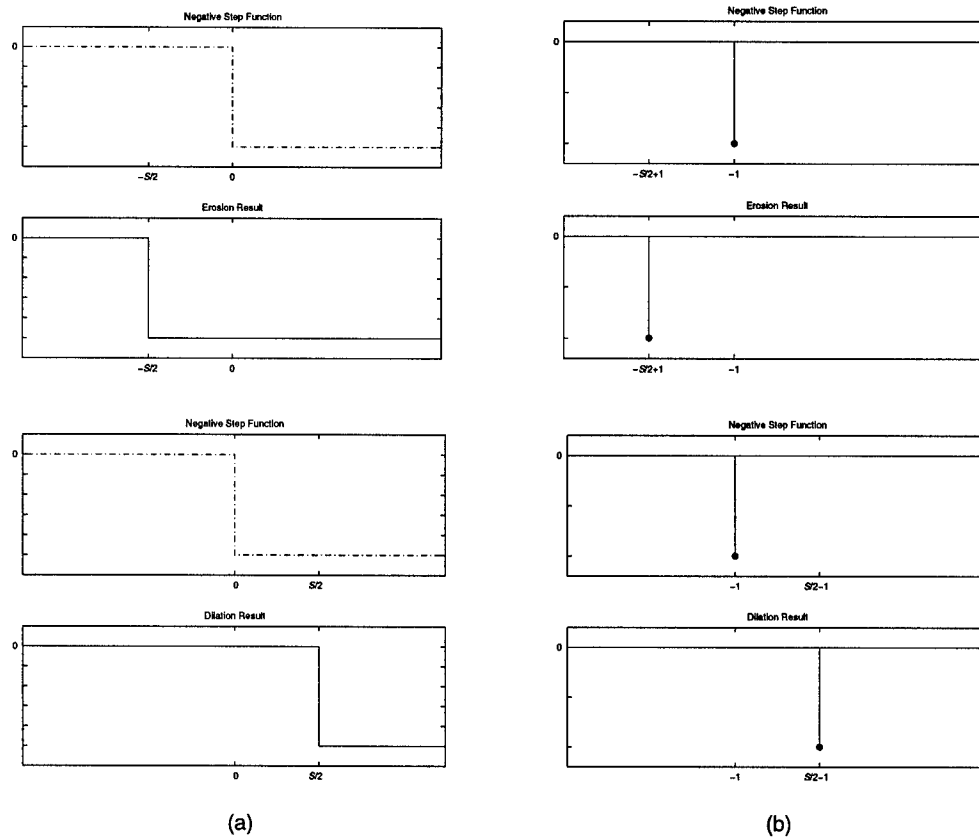


Figure 8. Effect of the fundamental morphological filters on a negative step function. The equivalent gradient representations are shown in (b).

Morphological filtering occurs as negative and positive gradients interact and remove each other. As displayed with the finite width edge model, eroding the function will eliminate it as the structuring element of the morphological filter becomes large. Results of filtering the finite width edge function using an erosion operation with several different structuring element sizes are shown in Figure 9, and equivalent gradient representations are shown in Figure 9b.

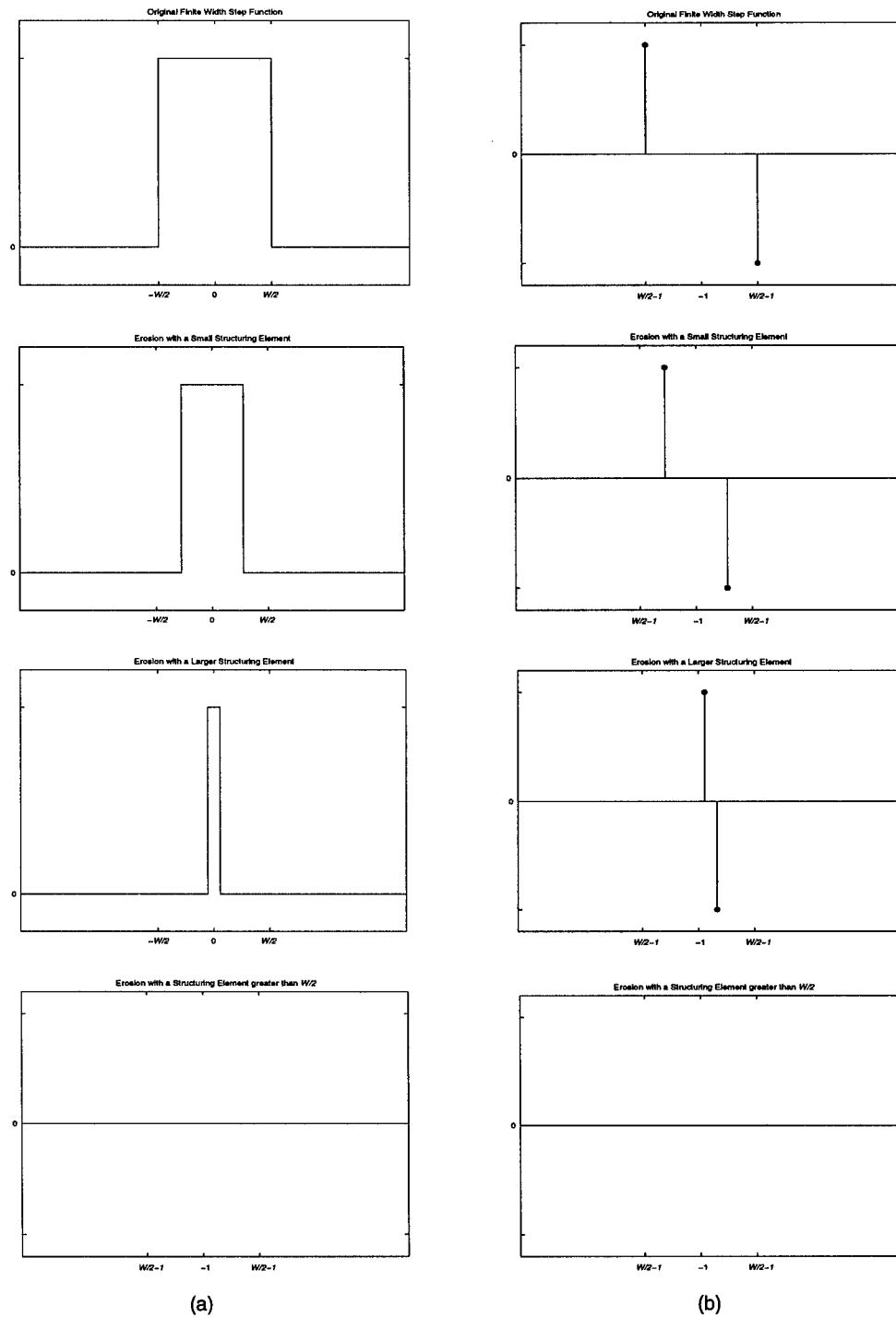


Figure 9. Eroding a finite width step function with structuring elements of different size. From top to bottom: the original finite width step function; the finite width step function eroded with a small structuring element; the finite width step function eroded by a larger structuring element; the finite width step function eroded with a structuring element larger than  $W/2$ . The equivalent gradient representations are shown in (b).

Analysis of the morphological operators becomes more difficult to visualize with an arbitrary signal. Relying on the eigenfunction properties of the step function, the sign dependent gradient representations, presented above, define the performance of the fundamental morphological filters with respect to edge preservation and movement. An erosion may be expressed as

$$\nabla \mathbf{I}_{t+\Delta t}(x) = \nabla^+ \mathbf{I}_t \left( x - \frac{\Delta t}{2} \right) + \nabla^- \mathbf{I}_t \left( x + \frac{\Delta t}{2} \right), \quad (10)$$

where  $t+\Delta t$  is the width of the structuring element,  $\nabla \mathbf{I}_t$  is the original image,  $\nabla^+$  is the maximum value of either the gradient or zero,  $\nabla^-$  is the minimum value of either the gradient or zero, and  $\Delta t$  is the time step. A dilation is realized by reversing the direction of gradient propagation and expressed as

$$\nabla \mathbf{I}_{t+\Delta t}(x) = \nabla^+ \mathbf{I}_t \left( x + \frac{\Delta t}{2} \right) + \nabla^- \mathbf{I}_t \left( x - \frac{\Delta t}{2} \right). \quad (11)$$

For discrete implementations of these fundamental morphological filters,  $\Delta t$  should be one.

Application of a single erosion or dilation operation removes regions dependent on intensity. It also results in edge movement. In the anisotropic diffusion equation, it is of interest to remove regions independent of intensity and without inducing edge movement. The sequential application of the fundamental filters can produce morphological systems realizing these goals. An *open* operation removes regions of low intensity, without introducing feature drift, and is implemented by first eroding a signal and then dilating the result. The *close* operation removes regions of high intensity, without inducing feature drift, and is implemented by dilating and eroding. Combinations of these higher level processes remove regions independent of intensity and without edge movement.

The purpose of this section is to show that incorporating morphological filters into the anisotropic diffusion equation creates a process that can remove features based solely on gradient or spatial scale and is, therefore, applicable to image pyramid construction. This section investigates the smoothing performance of the morphological filters on the step and finite width edge models. All analysis is accomplished using the previously defined gradient representations of the morphological filters and the anisotropic diffusion equation.

### Spatial Smoothing

The spatial smoothing of a morphological anisotropic diffusion system is displayed with the finite width step function. Considering a step function whose amplitude is greater than the smoothing threshold,  $\beta(K_{DC}, t)$ , traditional anisotropic diffusion expressions will preserve the gradients and maintain the feature. The goal of a scale aware diffusion process is to remove the region, independent of gradient magnitude.

In describing a morphological filter which, when incorporated into the diffusion coefficient, will smooth the finite width step function, it is necessary to define two properties of the morphological filtering operation: the direction of gradient propagation and the distance of gradient translation. The first characteristic, the direction of gradient propagation, is defined by the morphological filter type and denotes whether positive gradients are shifted to the right or left. (Negative gradients will be shifted in the opposite

direction.) The second characteristic, the distance of gradient translation, describes the spatial distance over which the gradients are moved and is defined by the solution time of the morphological system or, equivalently, the structuring element size.

Using a morphological filtering operation to smooth a finite width step function provides an initial description of morphological sequences suitable for incorporating into the diffusion equation, and filtering the finite width step function requires that the gradients of the edge model interact and remove each other. Remembering that the finite width step function is constructed with a positive gradient located at the origin and a negative gradient located at location  $W$ , where  $W$  is the width of the edge model, feature removal necessitates that gradients move towards each other. It also requires that positive gradient must move to its right while the negative gradient travels to its left.

Construction of the finite width step function defines the direction of gradient propagation necessary for filtering, and interaction between the two gradients of the edge model defines the distance of gradient translation. Removal of the finite width edge function will occur when the positive and negative gradients interact, and since the gradients move towards each other with equal speed, their interaction will occur at the center of the edge model. Specification of a morphological operation which allows the diffusion equation to smooth small objects requires a morphological filtering sequence which shifts positive gradients to the right a distance of  $W/2$ . This requirement is shown graphically in Figure 10, and an example morphological sequence satisfying these requirements is shown in (10), described by an erosion operation. Solved for solution times greater than  $W$ , the erosion operation will translate positive gradients to the right a distance of  $W/2$  and is a candidate filter for inclusion within a scale aware diffusion coefficient.

Scale aware diffusion expressions should remove regions of width  $W$ , and the erosion operation presented above satisfies this requirement. Scale aware diffusion expressions should also remove regions smaller than the spatial smoothing goals of the anisotropic diffusion equation, and analysis of smaller finite width edge models develops conditions necessary for their removal. Smoothing of a finite width step function with width  $n$  ( $0 < n \leq W$ ) still requires positive gradients be transported to the right, but only necessitates that they be translated over a distance of  $n/2$ . The requirements for removing all finite width step functions of

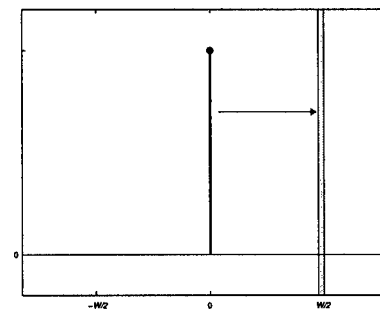


Figure 10. Necessary gradient movement for removing a finite width step function of width  $W$  with a morphological filter.

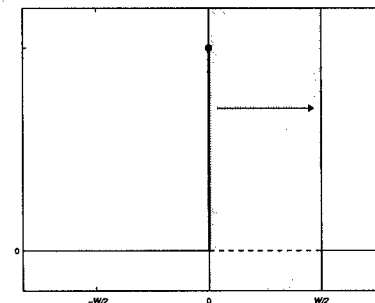


Figure 11. Necessary gradient movement for removing all finite width step functions of width less than or equal to  $W$  with a morphological filter.

width  $W$  or less are shown graphically in Figure 11. An example morphological sequence capable of satisfying these conditions is the erosion operation solved solution times greater than  $W$  - the identical filtering operation derived for smoothing the larger finite width step functions.

While study of the smaller finite width edge model did not refine the requirements of candidate morphological sequences, analysis of the negative finite width step function further constrains the construction of the morphological filter and motivates the need for a more complex filter sequence. The negative finite width step function is given as

$$u_{fw}(x) = \alpha u(n) - \alpha u(0), \quad 0 < n \leq W \quad (12)$$

and should also be removed by the scale aware diffusion equation. Smoothing of this function introduces different requirements on the direction of gradient propagation within the morphological filter sequence, as the positive gradient of this edge model is located to the right of the negative gradient. (The positive gradient of the original finite width step function was located to the left of the negative gradient.) Filtering necessitates that the gradients move towards one another, and for the negative finite width step function requires that the positive gradient travel to its left and the negative gradient travel to its right. Feature removal will occur when the gradients interact at the center of the edge model,  $n/2$ , and Figure 12 summarizes the requirements for removing all of the negative finite width edge models. A candidate morphological filter providing this smoothing is the dilation operation presented in (11) and solved for solution times greater than  $W$ .

Consolidating the requirements for removing both finite width step functions concludes this subsection and defines the class of morphological filters suitable for providing a diffusion coefficient with the capability of identifying and smoothing regions of small spatial size. It has been shown that removing a positive finite width step function requires a morphological filter sequence which translates positive gradients to the right and that smoothing a negative finite width step function requires a filtering operation capable of translating positive gradients to the left. In both smoothing examples, the original gradients must be transported over a distance equal to half the spatial smoothing goals of the anisotropic diffusion equation,  $W/2$ . These requirements are summarized graphically in Figure 13.

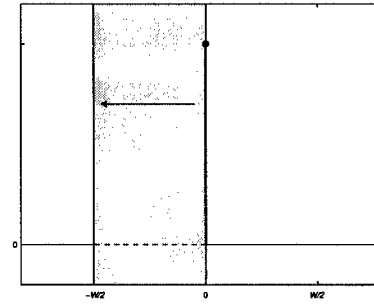


Figure 12. Necessary gradient movement for removing all negative finite width step function of width less than or equal to  $W$  with a morphological filter.

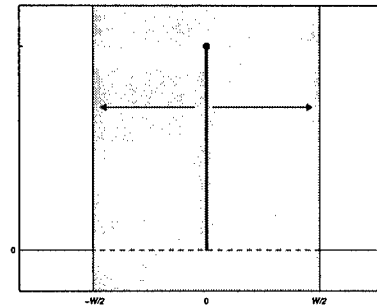


Figure 13. Necessary gradient movement for removing all negative and positive finite width step functions of width less than or equal to  $W$  with a morphological filter.

While the single application of an erosion or dilation operation will not generate the necessary smoothing performance for filtering both edge models, the sequential concatenation of these fundamental morphological operators will produce a filter capable of satisfying these requirements, transporting gradients throughout the desired regions. Many morphological operations could be constructed, and an example morphological filter sequence, which produces the necessary gradient movement, is shown graphically in Figure 14. The sequence is realized by dilating the signal with solution time  $W+1$  and then eroding the result with solution time  $2(W+1)$ . This initially moves the positive gradient over the region to the left, removing positive finite width edge models, and then translates the gradient back to the origin and through the region to the right, removing negative finite width edge models.

Analysis of the finite width step functions defines a class of morphological operators which remove features of small scale and whose incorporation into a diffusion coefficient would allow the anisotropic diffusion expression to smooth these small scale regions. The purpose of the next subsection is to define a class of morphological filters which allows the anisotropic diffusion equation to smooth regions of low contrast. After deriving these smoothing conditions, the section will conclude that incorporating certain morphological filter sequences into a diffusive process develops a smoothing operation capable of simultaneously removing objects of small spatial size while smoothing gradients of small scale.

### Gradient Smoothing

Filtering of small scale objects necessitates that positive and negative gradients interact, motivating morphological operators to be described through the regions over which a gradient must travel. Smoothing small contrast regions requires that significant gradient magnitudes not interfere in diffusion. In describing morphological filters that do not inhibit the smoothing of low contrast areas, filter types are defined by the net distance of gradient movement and are unconcerned with the specific path a gradient undertakes. Consider the sequence of two step functions shown in Figure 15. Application of morphological operators to this sequence will never change the structure of the signal, as both gradients will travel in the same direction and at the same speed. Morphological operators applied to monotonic regions can only introduce delay.

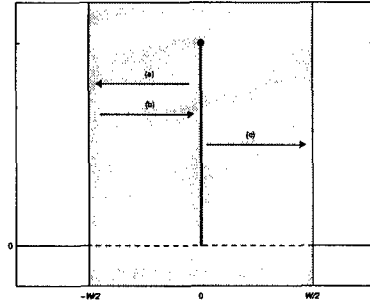


Figure 14. Necessary gradient movement for removing all negative and positive finite width step functions of width less than or equal to  $W$ . The morphological sequence consists of: (a) dilation, (b) erosion, and (c) erosion.

Defining the height of the first step function to have small magnitude ( $\gamma \cong 0$ ) and the height of the second step function to be significant ( $\alpha \geq \beta(K_{DC}, t)$ ), traditional anisotropic diffusion expressions will maintain the larger edge function while smoothing the smaller step. An ideal scale aware diffusion expression attempts to reproduce the smoothing properties of traditional anisotropic diffusion in the absence of small spatial features. Incorporating morphological operators into the diffusion coefficient requires that larger gradients not be translated to the positions occupied by the smaller step function. The only spatial location guaranteed not to contain a smaller gradient is at the original location of the larger gradient. Therefore, the simple criterion that must be satisfied by a morphological filter is that the morphological sequences produce a net translation of zero.

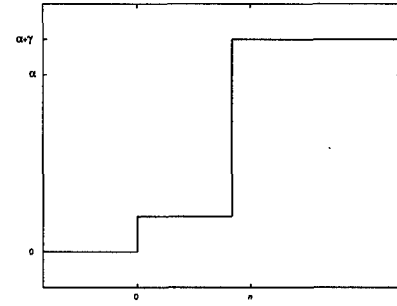
This condition on the construction of morphological filters that, when incorporated into the diffusion coefficient, allows the smoothing of regions of low contrast completes this section. Analysis of finite width edge models developed criteria on the morphological operators for the identification of small scale objects and required specific regions through which gradients must travel. Analysis of the second edge sequence introduced no further constraints on the path of gradient movement, but only defined the morphological filters to have a net gradient translation of zero. Morphological operators, unlike the linear filters, exist which are capable of simultaneously satisfying these conditions. As an example, the morphological open-close filter is constructed from an erosion-dilation-dilation-erosion sequence and propagates the gradients throughout the necessary region of influence while introducing a net translation of zero.

Introducing morphological operators into the diffusion coefficient is straightforward. Using the coefficient presented in (5) as an example, a possible realization of a morphological scale aware diffusion coefficient is

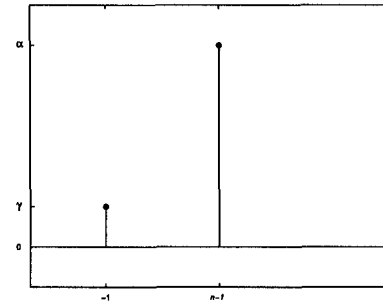
$$c(\nabla \mathbf{I}) = e^{-\left(\frac{\|\nabla((\mathbf{I} \circ \mathbf{M}) \bullet \mathbf{M})\|}{k}\right)^2}, \quad (13)$$

where  $(\mathbf{I} \circ \mathbf{M}) \bullet \mathbf{M}$  is the open-close filtering of the signal  $\mathbf{I}$  with structuring element  $\mathbf{M}$ .

To display the effectiveness of a morphological diffusion coefficient, simulations were conducted with three classes of the anisotropic diffusion coefficient. The first class was dependent solely on local gradient information and represented by the traditional diffusion coefficient expression, given as



(a)



(b)

Figure 15. A sequence of two step functions with arbitrary height. The equivalent gradient representation is shown in (b).

$$c(\nabla I) = e^{-\left(\frac{\|\nabla I\|}{k}\right)^2}, \quad (14)$$

where  $I$  is the original image and  $k$  is the gradient threshold. The second and third classes employed scale-modified definitions for the gradient. The second class used a linear filter in performing the gradient calculation, while the third class used a nonlinear morphological filter. These classes were represented by diffusion coefficients described by

$$c(\nabla I) = e^{-\left(\frac{\|G_\sigma * \nabla I\|}{k}\right)^2}, \quad (15)$$

where  $G_\sigma$  is a Gaussian kernel with standard deviation  $\sigma$ , and

$$c(\nabla I) = e^{-\left(\frac{\|\nabla((I \circ M) \bullet M)\|}{k}\right)^2}, \quad (16)$$

where  $M$  is a morphological structuring element and  $(I \circ M) \bullet M$  is the image  $I$  filtered with an open-close filter.

Equivalent scale parameters for the linear and morphological filters in the scale aware diffusion coefficients were chosen to provide information removal of similar scale, and both were defined by satisfying conditions necessary for decimating by a factor of three. The Gaussian kernel used in the second coefficient class was defined to have a standard deviation of  $6/\pi$ , as suggested to approximately satisfy Shannon's sampling theorem in (Burt, 1981). Similarly, the morphological kernel used in the third coefficient class was defined to be a square structuring element of width five, as suggested to approximately satisfy the Homotopy Preserving Critical Sampling Theorem in (Florencio, 1994).

Producing a qualitative evaluation of the three processes, the anisotropic diffusion equation was applied to infrared imagery of a T-72 tank. Results are shown in Figure 16. As may be observed, the diffusion equation based solely on local gradient information is unable to remove channel noise and fine detail. The second coefficient class, using a linear filter within its gradient calculation, removes these small features, but at the expense of introducing edge movement and feature drift. (Notice the excessive smoothing of the tank edges.) The morphological anisotropic diffusion algorithm is capable of overcoming both deficiencies, removing the noise while maintaining edges. Quantitative studies are provided in the next section.

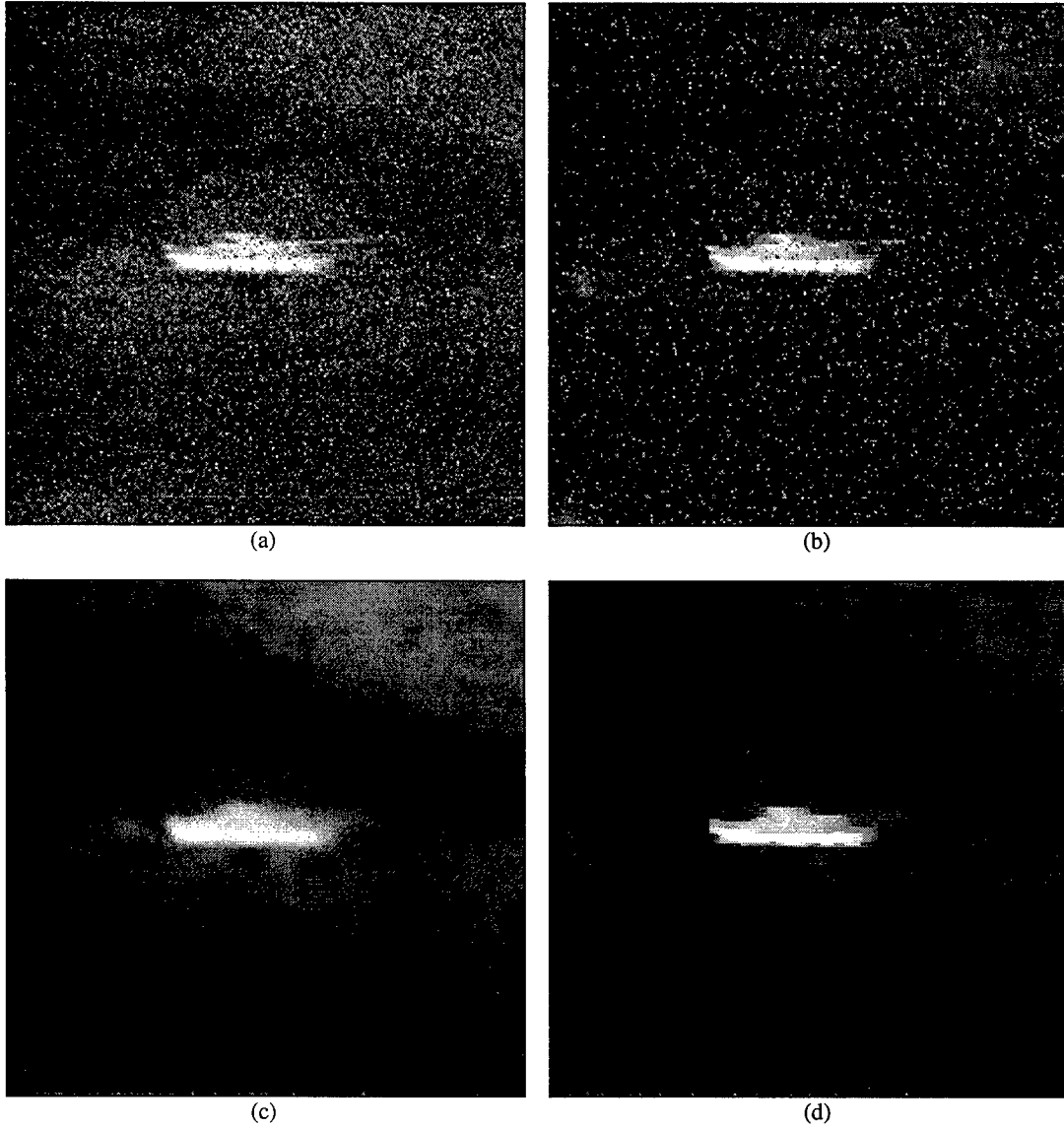


Figure 16. Three classes of anisotropic diffusion applied to the T-72 image: (a) original image; (b) results obtained using original anisotropic diffusion; (c) results obtained using traditional scale aware anisotropic diffusion; (d) results obtained using morphological anisotropic diffusion (16). The morphological anisotropic diffusion process removes extraneous information introduced by various environmental factors. Edges are also preserved, and the infrared imagery is enhanced.

## Morphological Diffusion Simulation

To display the effectiveness of the new morphological diffusion coefficient, simulations were conducted with three classes of the anisotropic diffusion coefficient. The first class was dependent solely on local gradient information and represented by the traditional diffusion coefficient expression, given as

$$c(\nabla \mathbf{I}) = e^{-\left(\frac{\|\nabla \mathbf{I}\|}{k}\right)^2}, \quad (17)$$

where  $\mathbf{I}$  is the original image and  $k$  is the gradient threshold. The second and third classes employed scale-modified definitions for the gradient. The second class used a linear filter in performing the gradient calculation, while the third class used a nonlinear morphological filter. These classes were represented by diffusion coefficients described by

$$c(\nabla \mathbf{I}) = e^{-\left(\frac{\|\nabla(\mathbf{G}_\sigma * \mathbf{I})\|}{k}\right)^2}, \quad (18)$$

where  $\mathbf{G}_\sigma$  is a Gaussian kernel with standard deviation  $\sigma$ , and

$$c(\nabla \mathbf{I}) = e^{-\left(\frac{\|\nabla((\mathbf{I} \circ \mathbf{M}) \bullet \mathbf{M})\|}{k}\right)^2}, \quad (19)$$

where  $\mathbf{M}$  is a morphological structuring element and  $(\mathbf{I} \circ \mathbf{M}) \bullet \mathbf{M}$  is the image  $\mathbf{I}$  filtered with an open-close filter.

Equivalent scale parameters for the linear and morphological filters in the scale aware diffusion coefficients were chosen to provide information removal of similar scale, and both were defined by satisfying conditions necessary for subsampling a filtered representations by a factor of three. The Gaussian kernel used in the second coefficient class was defined to have a standard deviation of  $6/\pi$ , as suggested to approximately satisfy Shannon's sampling theorem in (Burt, 1981). Similarly, the morphological kernel used in the third coefficient class was defined to be a square structuring element of width five, as suggested to satisfy the Homotopy Preserving Critical Sampling Theorem in (Florencio, 1994).

Producing a qualitative evaluation of the three processes, the anisotropic diffusion equation was applied to synthetic imagery corrupted by 40% salt and pepper noise. Results are shown in Figure 17. As can be seen, the diffusion equation based solely on local gradient information is unable to remove impulsive noise, while both spatially enlarged coefficients are capable of smoothing these small, high contrast objects and maintaining large scale edges. Results for coefficient classes two and three are visually similar, although closer inspection will show that the third class, the nonlinear morphological method, provides a slight improvement in feature definition.

A second qualitative example of the three anisotropic diffusion processes was attained by applying the smoothing operations to the cameraman image. These results are similar to those achieved with the previous synthetic imagery, and they are presented in Figure 18. Again, the first coefficient class, using a traditional gradient calculation, is unable to remove fine detail, as evident by the existence of the small objects present on the ground. The second coefficient class, using a linear filter within its gradient calculation, removes these

small features, but at the expense of introducing edge movement and feature drift. (Notice the excessive smoothing of the camera and the movement of the elbow.) The new morphological anisotropic diffusion algorithm is capable of overcoming both deficiencies, removing small objects while maintaining edge locality.

While qualitative comparison of the three methods of anisotropic diffusion begins to distinguish the smoothing properties of the morphological diffusion coefficient, a quantitative comparison of their edge detection accuracy displays the advantages of the new diffusion expression. In determining the edge detection capabilities of the three variants of anisotropic diffusion, synthetic imagery was again corrupted by 40% salt and pepper noise. These images were then smoothed; and at each solution time, edges were identified and compared with known edge locations. Recognizing edges in the filtered imagery was accomplished with the use of a simple gradient based edge detector, well motivated by the smoothing properties of the anisotropic diffusion equation, and the threshold of the edge detector was defined to be equal to the gradient threshold of the diffusion coefficient,  $k$ .

Experimental comparison of edge detection performance was calculated using two quantitative metrics. The first, Pratt's edge quality measurement, is defined as

$$F = \frac{\sum_{i=1}^{I_A} \frac{1}{1 + \alpha(d(i)^2)}}{\max\{I_A, I_I\}} \quad (20)$$

where  $I_A$  is the number of edge points detected in the filtered image result,  $I_I$  is the number of edge points existing in the original, noise free imagery,  $d(i)$  is the Euclidean distance between an edge location in the original image and the nearest detected edge, and  $\alpha$  is a scaling constant, set to the suggested value of 1/9 (Pratt, 1978). Perfect recovery of all edge information in the original image results in an edge quality measurement of one ( $F=1$ ); poor edge localization lowers the value.

The second measurement group contains two more tangible representations of the candidate filter performance, and the first measurement is defined to be the percentage of original edge points successfully identified by the edge detection process. Correctly recovering all edges in the initial image results in a 100% identification percentage, not detecting a feature at its original location lowers the identification measurement. The second measurement describes the ability of the edge detector to identify edges without detecting false edge locations. Expanding on the previous measurement, edge features which are not recognized and image locations which are erroneously classified as features are calculated. Perfect recovery of the original image results in an identification measurement of 100%, incorrect identification of any image location lowers the measurement.

The results of the numerical experiment are presented in Figure 19. It may be seen that the linear coefficient initially outperforms the other diffusion variants in the edge quality measurement, but produces the poorest identification percentage. As solution time increases, the introduction of edge localization errors by the linear filter becomes more evident and is displayed by the rapid decrease in matched features. Specifically, at solution time three, the linear coefficient is unable to correctly identify the location of a single edge. The morphological anisotropic diffusion method provides significant performance improvement,

able to identify over 70% of the original edges and attain a solution quality measurement of 0.95.

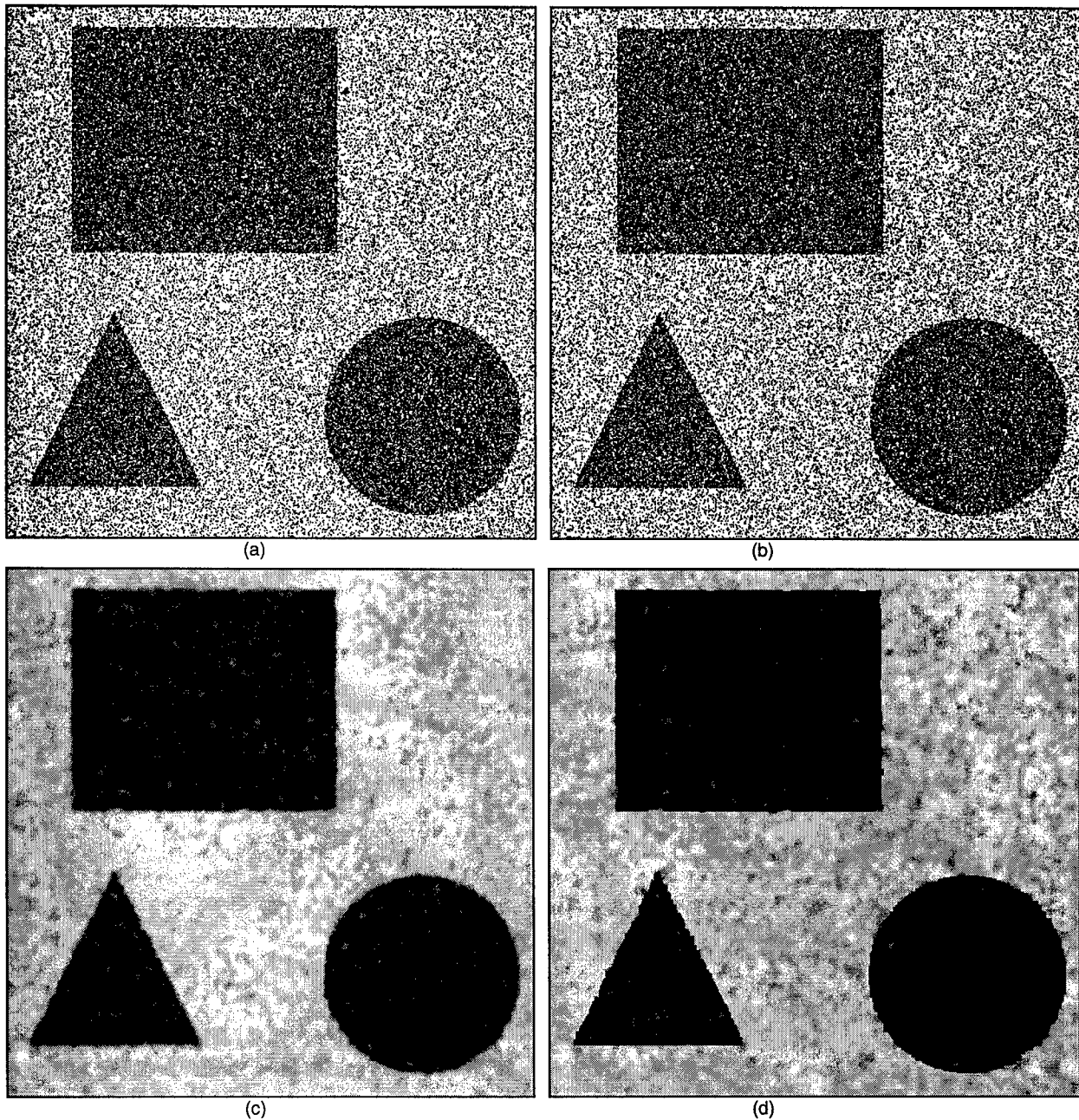


Figure 17. Three classes of anisotropic diffusion applied to synthetic imagery: (a) original image corrupted with 40% salt and pepper noise; (b) results obtained using original anisotropic diffusion; (c) results obtained using traditional scale aware anisotropic diffusion; (d) results obtained using morphological anisotropic diffusion (19). The gradient threshold,  $k=40$ , and the solution time,  $t=3$ .



(a)



(b)



(c)



(d)

Figure 18. Three classes of anisotropic diffusion applied to the cameraman image: (a) original image; (b) results obtained using original anisotropic diffusion; (c) results obtained using traditional scale aware anisotropic diffusion; (d) results obtained using morphological anisotropic diffusion (19). The gradient threshold,  $k=10$ , and the solution time,  $t=20$ .

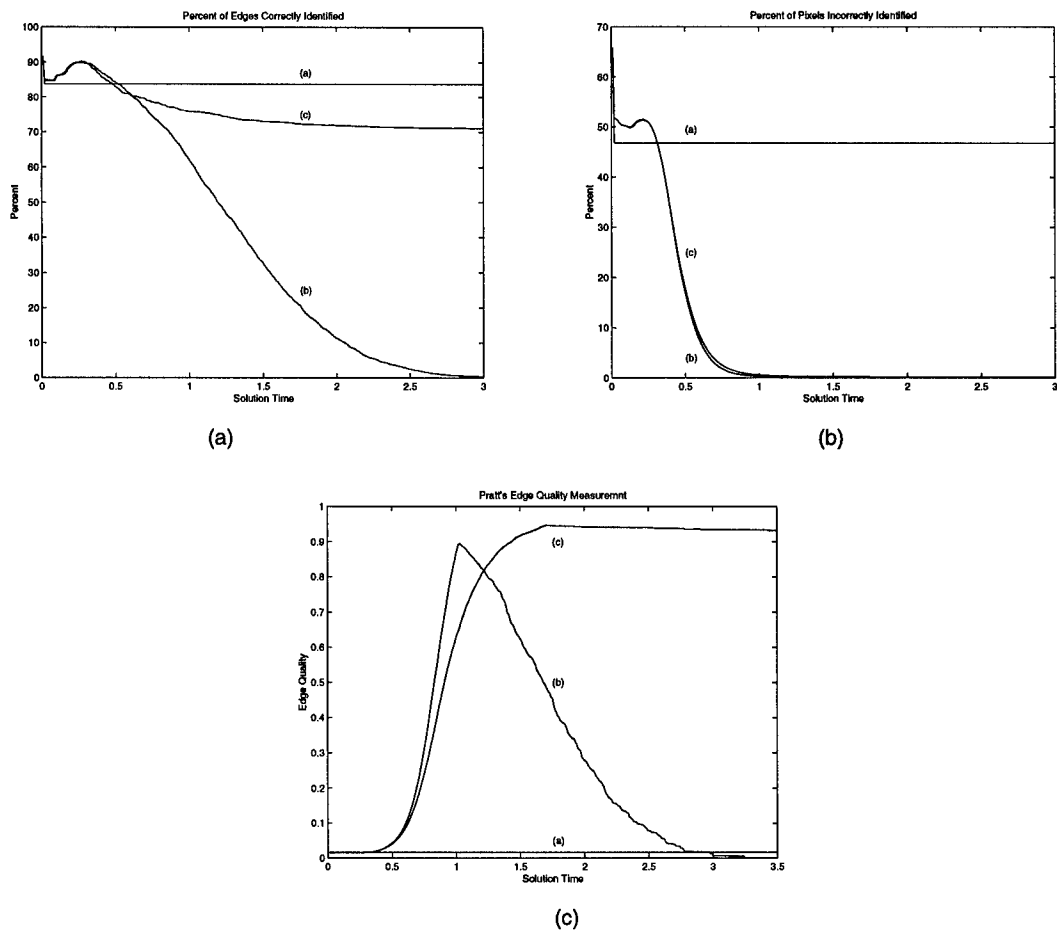


Figure 19. Three quantitative measurements of edge detection performance: (a) percent of edges correctly identified; (b) percent of image pixels incorrectly classified by the edge detector; (c) results of Pratt's edge quality measurement.

### Anisotropic Diffusion Pyramids

Construction of an anisotropic diffusion pyramid (ADP) requires selecting the scale parameters of the anisotropic diffusion equation. These parameters should allow the filtered result to be sampled without loss of information. In applying a pyramid structure to the IRST problem, sampling is not viewed within the context of reconstruction but rather within the context of spatial causality. Spatial causality describes a cause and effect relationship between scale representations. In this section, the construction of image pyramids that utilize the anisotropic diffusion equation as the scale generating operator will be considered. Throughout this discussion, the anisotropic diffusion equation will be analyzed using continuous input signals and treated as a piece-wise linear operator. Approximations of the continuous diffusion expressions with discrete difference equations will also be considered.

Constructing an image pyramid with a fixed scale filter is possible only if the scale generating function also produces a signal suitable for sampling. The diffusion equation must therefore smooth all internal features and reduce their gradients below the gradient threshold. The gradient of a sampled signal representation is proportional to the gradient of the original representation by the sample factor  $S$ . A guarantee that small spatial features will be removed in coarser scale depictions is given by smoothing all internal features such that all gradients are less than  $k/S$ , where  $k$  is the gradient threshold of the diffusion coefficient.

To derive a solution time that satisfies this requirement, we consider a single finite width step function of small spatial size and height  $\alpha$ . The filtered representation of the finite width step function is then expressed as

$$\alpha \cdot \left( \frac{1}{\sqrt{4\pi t}} e^{\frac{-x^2}{4t}} \right), \quad 0 \leq t \leq \infty \quad (21)$$

where  $\alpha$  is the magnitude of the impulse and  $t$  is the solution time of the scale aware anisotropic diffusion expression. Examining the derivative of a filtered impulse signal, smoothing criteria becomes apparent. Edges will be removed from subsampled representations if their gradients are less than  $k/S$ . Solution times assuring removal must satisfy

$$\left| \frac{\partial}{\partial x} \left( \alpha \cdot \frac{1}{\sqrt{4\pi t}} e^{\frac{-x^2}{4t}} \right) \right| < \frac{k}{S}, \quad 0 \leq t \leq \infty \quad (22)$$

where  $\alpha$  is the magnitude (maximum intensity) of the impulse,  $k$  is the gradient threshold of the edge detection system,  $t$  is the solution time of the scale aware anisotropic diffusion expression, and  $S$  is the sample factor. Solving this equation, the value of the smoothing parameter,  $t_s$ , is defined as

$$t_s > \frac{\alpha \cdot S}{k\sqrt{8\pi}} e^{\frac{1}{2}}, \quad (23)$$

where  $S$  is the sample factor. For an iterative approximation of the continuous anisotropic diffusion equation, the number of diffusion iterations necessary for suitably filtering a signal

would be approximately  $t_s \cdot \Delta t$ , where  $\Delta t$  is the time step used in the discrete realization of the diffusion equation (Niessen, 1994).

The morphological diffusion process contains a second description of scale. This additional filter parameter specifies the scope of the diffusion coefficient calculation and defines the spatial size of objects that are determined to be “small” and subsequently removed. Borrowing from the study of morphology, an image region may be viewed as a set and the region over which a coefficient is calculated defined. The diffusion system must preserve an object’s *homotopy* across sampling domains, where homotopy is simply a one-to-one mapping of objects. Similar to the frequency based sampling strategy, homotopy will only be guaranteed if all sets are removed that are spatially unsupported by the new sample domain. This requires the identification and smoothing of all objects smaller than the sample grid, so that after sampling they will not exist. Using morphological operators within a diffusion coefficient, the spatial scale of diffusion is defined by the structuring element size of the morphological operators. Ensuring the identification and removal of all objects smaller than the sample grid, the morphological operators utilize structuring elements with diameter greater than the sample factor. For constructing an anisotropic diffusion coefficient, the structuring elements used by the morphological filters must have diameter of  $\sqrt{2} \cdot S$ , where  $S$  is the sample factor (Florencio, 1994).

With this result, an ADP may be constructed by successively filtering and subsampling previous resolution representations. Using the traditional discrete approximation of the partial differential equation presented in (3), construction of coarser resolution images within an ADP may be expressed as

$$\mathbf{I}_L = \left[ \mathbf{I}_{L-1} + \Delta t \sum_{m=1}^{t_s(\Delta t)} (c_N \nabla_N + c_S \nabla_S + c_E \nabla_E + c_W \nabla_W)_m \right]_{\downarrow S}, \quad (24)$$

where  $t_s$  is the solution time ensuring spatial causality from (23),  $\downarrow S$  denotes subsampling by a factor of  $S$ ,  $\Delta t$  is the time step, and  $c$  is the scale aware diffusion coefficient. For the generation of an ADP, the original image is filtered with anisotropic diffusion until the appropriate solution time. This intermediate representation is then subsampled, and the result is the first pyramid level. (The original image is defined as the zero level within an image pyramid.) Higher levels of the multi-scale structure are computed by filtering and subsampling previous resolution representations.

### **Multi-scale Tracking**

The advantage of using a multi-resolution IRST technique is embedded in the utilization of coarse scene representations for the initial identification of a target. Coarser scene information is created by successively filtering and subsampling the original image, and its use allows initial object identification to query information absent of noise and represented at reduced sample densities. Maximizing the benefits of these coarse-to-fine search procedures is accomplished by initially identifying an object at the coarsest resolution possible. In a multi-scale procedure, this level is defined to be the *root level* of the search.

The root level of an object simply defines the coarsest resolution representation in which the object will be identifiable. The procedure used for selecting the root level of a coarse-to-fine search is best illustrated with an example. Consider the image presented in Figure 20, consisting of two aircraft. In determining the root level necessary for the identification of the smaller plane, two scene measurements must be estimated. The first measurement is the minor axis length of the largest target feature. This measurement describes the pyramid level at which the target will be removed, and in this example, the distance 18 pixels and is denoted graphically with the variable  $y$ . The second scene estimate to be acquired is the minimum distance between the smallest object and other large scale features. This measurement describes the pyramid level in which the two objects will merge. In the figure, the distance is represented with the variable  $z$  and measured to be 55 pixels. After identifying the two scene measurements, computation of the root level is accomplished by the procedure outlined below. In the example, the root level of the smaller aircraft is found to be the third level of the image pyramid.

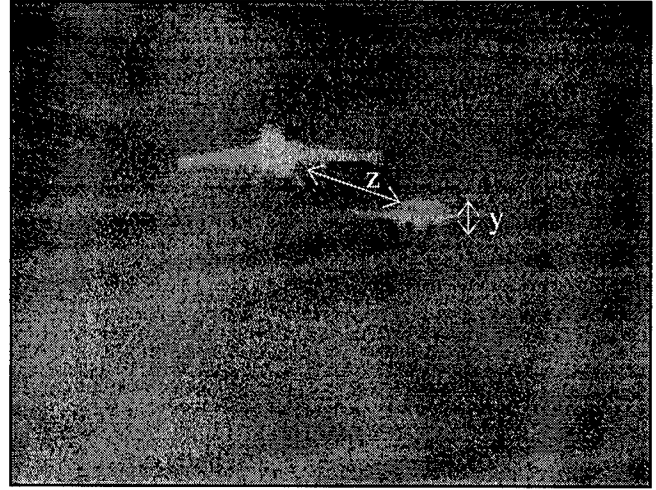


Figure 20. Designing a coarse to fine search that is capable of identifying the smaller aircraft in the infrared image. Root level selection is dependent on two variables, the minor axis of the target's largest feature ( $y$ ) and the minimum distance between the target and other objects of similar or greater scale ( $z$ ).

The first step in deriving root level definitions is to model the arbitrary object as a collection of smaller convex features. As an example, the jet aircraft in Figure 20 may be modeled as a composition of three smaller features: the fuselage, wing, and landing gear. Expressing an arbitrary object,  $O$ , as the union of a set of smaller convex sets, the object may be described as

$$O = \bigcup_{m=1}^M o_m \text{ where } o_i \cap o_j = \emptyset \text{ for } \forall i, j \leq M, \quad (25)$$

where  $O$  is the object of interest and  $o_m$  are individual features. Considering all targets as a collection of convex features, the derivation of an object's root level is straightforward.

Defining the root level of an object with respect to its own internal composition is the first selection criteria considered in this subsection, and it necessitates defining the coarsest pyramid level that contains the object. The anisotropic diffusion expression removes all regions of small spatial scale, and for the complete removal of a target, the anisotropic diffusion process must remove all features of the target. These features will be removed according to their spatial size, with smaller features being removed before larger ones, and the selection of pyramid levels which contain the target require the representation to contain the largest target feature.

Subsampling a target feature reduces its spatial dimension, and in an anisotropic diffusion pyramid, spatial measurements of large objects, before and after sampling, are related by the proportionality factor  $S$ , where  $S$  is the sample factor used for pyramid construction. Subsampled objects have dimensions that are  $S^{-1}$  the size of the original object, and the size of a large object at pyramid level  $L$  may be expressed as

$$y_L = \frac{y}{S^L}, \quad (26)$$

where  $y$  is the original measurement and  $y_L$  is the equivalent spatial dimension in the subsampled domain.

Continuously filtering and subsampling an object should eventually result in its removal, and within an anisotropic diffusion pyramid, a feature is defined to be removed when its spatial size is smaller than the sample grid ( $y_L < S$ ). The largest feature of the target will disappear in the construction of pyramid level  $L+1$ , when

$$S > \frac{y}{S^L}, \quad (27)$$

where  $y$  is the smallest spatial dimension of the feature (the minor axis),  $S$  is the sample factor, and  $L$  is the previous pyramid level.

Rearranging this equation produces the first definition of the root level of an object. The coarsest pyramid level in which a target will exist may be expressed as

$$L_R = \max\{L: L < \log_S |y| - 1\}, \quad (28)$$

where  $L_R$  is the root level defined by the internal characteristics of an object,  $S$  is the sample factor used in pyramid construction, and  $y$  is the minor axis of the largest convex feature of the target.

While the root level may be described by its internal composition, a complete definition must consider the content and construction of its environment. A target may also elude identification when multiple objects merge, as the search procedure will no longer be capable of resolving either individual object. Describing the separation distance between the target and the second object with the distance  $z$ , the two objects will merge in the construction of pyramid level  $L+1$ , when

$$S > \frac{z}{S^L}, \quad (29)$$

where  $z$  is the minimum distance between the two objects,  $S$  is the sample factor used in the construction of the pyramid, and  $L$  is the previous pyramid level. Rearranging (29) presents the second description of the root level of a target, expressed as

$$L_R = \max\{L: L < \log_S |z| - 1\}, \quad (30)$$

where  $L_R$  is the root level defined by the external characteristics of a scene.

The definition of the root level used in a coarse to fine search is determined both by an object's internal and external characteristics, and it may now be defined as

$$L_R = \max\{L: L < \log_s |d| - 1\}, \quad (31)$$

where

$$d = \min\{y, z\}. \quad (32)$$

With the selection of the root level, a coarse-to-fine search procedure may be initialized to maximize the efficiency and structure of the ADP. These search methods identify an object within coarse resolution representations and then use these results to constrain higher resolution inspections. The practical realization of this search procedure begins by identifying the target within the root level. In the target tracking system used for simulation, identification utilizes binary edge maps of the candidate target and the current scene. Edge maps are constructed by thresholding the gradient of the image. An example of a multi-scale edge template is presented in Figure 21. Computing a binary exclusive-OR between coarse scale template information and the scene facilitates locating an object in the root level of the image pyramid. This operation may be described as



Figure 21. A multi-resolution template for a semi truck. Coarser template representations are used to search coarse scene descriptions within the pyramid.

$$Match(i, j) = \sum_x \sum_y \mathbf{T}_{L_R}(x, y) \oplus \mathbf{I}_{L_R}(x + i, y + j) \quad (33)$$

where  $\mathbf{T}_{L_R}$  is the template representation at the root level,  $\mathbf{I}_{L_R}$  is the scene representation at the root level, and  $\oplus$  denotes an exclusive-OR operation. Higher values for the binary template match correspond to higher similarity between template and scene, and in the simulation results to follow, the highest match is defined to correspond to the target.

Having identified the best match between scene and template at location  $(i, j)$ , the goal of a multi-scale IRST procedure is to use these results to guide and refine progressively higher resolution inspections. ADPs were designed to maintain spatial causality, ensuring that a features that exists at location  $(i, j)$  in a coarse resolution representation will exist within the region  $(S \cdot (i \pm 1/2), S \cdot (j \pm 1/2))$  in higher resolution depictions, where  $S$  is the sample factor used in the construction of the image pyramid (Burt *et al.*, 1981). Using this relationship, the realization of a coarse-to-fine search procedure is straightforward. Target identification begins by locating the best match between edge template and scene, using (33). Higher resolution information is then queried but only at four possible object locations. The identification results attained from inspecting the level  $L_{R-1}$  are then used to constrain the search of the next level,  $L_{R-2}$ , and the procedure terminates after finding the target in the original image,  $L_0$ .

## Experimental Results

Experimental results attest to the solution quality of the ADP. Here, three "real world" image sequences are processed. Comparisons are presented between two IRST approaches. The first method utilizes the multi-scale, multi-resolution ADP tracking mechanism. The second method is a traditional, single resolution, tracking algorithm. In the following simulations, solution quality will be described using two metrics: the measurement error between an object's identified location and the ground truth location and the computational requirements of the search routine. The results will illustrate the strengths of the ADP approach, specifically showing that it is a more robust and efficient solution to the search and track problem than signal resolution methods.

### Jet Sequence

The first image sequence used in the solution quality simulations consisted of 25 frames of a jet airplane in flight. Each original image, and the base of the corresponding pyramid, had a resolution of 256x256 pixels, and all pixels were capable of representing 256 intensity levels. The ADPs were constructed with a 1 of 2 uniform sampling scheme (along each row and column), a gradient threshold  $k = 15$ , and  $\Delta t = 1/4$ . Figure 22 displays the pyramid constructed for the first frame of the sequence. Implementing a multi-resolution search for the identification of the jet aircraft requires the definition of the root level, and the root level was defined to be the third resolution representation above the original image.

Applying the coarse-to-fine search techniques to the target identification problem and using the third level of the pyramid as the root level, object tracking tasks were performed using the edge-based template matching routine. The result was a significant increase in computational efficiency between the single resolution and multi-scale techniques. A single resolution match required approximately 172 seconds per frame on a Sun Ultra 1/170, while the multi-resolution approach required approximately 8 seconds per frame, including the

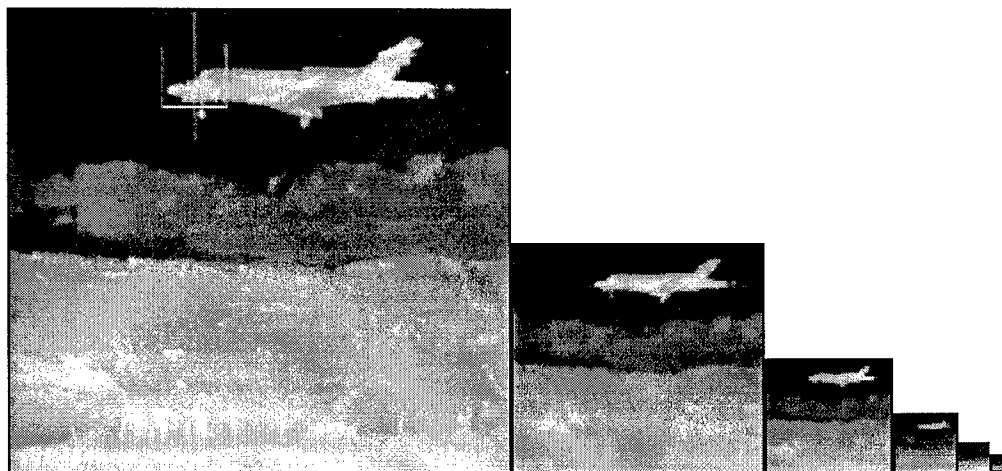


Figure 22. An infrared image of a jet airplane in flight and its corresponding ADP. The root level is the coarsest resolution that contains the aircraft. Edge features belonging to the aircraft are last found in the fourth largest image of the pyramid (the third level).

construction costs of the ADP. The effect was a system performance improvement of 21 times tradition single-resolution methods.

Besides providing computation efficiency, multi-resolution techniques also increase system robustness. Using the same binary template matching routine, pixel localization errors were computed for both single and multi-resolution trials. These results are summarized in Figure 23, where the localization errors are expressed as the Euclidean distance between identified object locations and ground truth. For the first 14 frames of the sequence, the algorithms produce similar measurements. In the final 11 images, the multi-scale search was able to locate the target while the single-resolution method was not. This displays the inability of the single-resolution method to accommodate slight changes between images and the template. The ADP is more resilient, taking advantage of the high similarity between coarse scale descriptions within the image sequence.

To further display the robustness of the ADP, simulations were performed on the same sequence of images, corrupted by Gaussian distributed noise. The mean-square signal to noise ratio of the test images was 15.72. As can be seen from the identification results presented in Figure 24, the pixel localization error of the multi-scale technique increased in the presence of the additive noise, but the coarse-to-fine search method was still capable of providing acceptable estimates of the object location. Conversely, the single resolution method was unable to reliably determine the location of the target during any frame of the sequence. The ability to find an object in high clutter allowed the multi-scale search and track system to provide a smaller pixel localization error, with a mean error of 3.69 pixels compared to 195.29 pixels of the single resolution system.

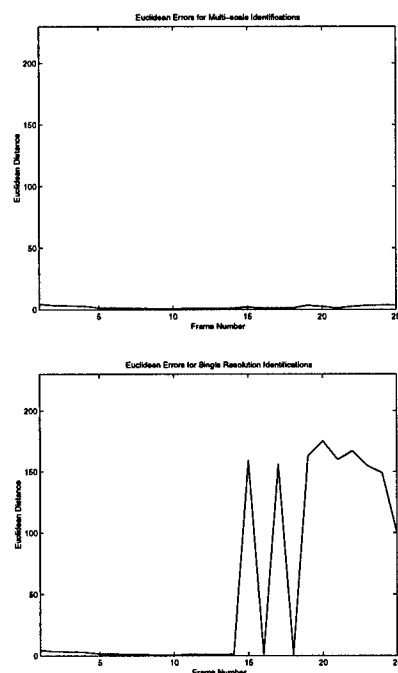


Figure 23. Localization errors for the jet sequence. Errors were calculated for both the single resolution and multi-resolution identification procedures by computing the Euclidean distance between identified target locations and ground truth. The two algorithm produce equivalent results for the first 14 frames of the sequence. In the later frames of the sequence, the single resolution technique does not reliably identify the target.

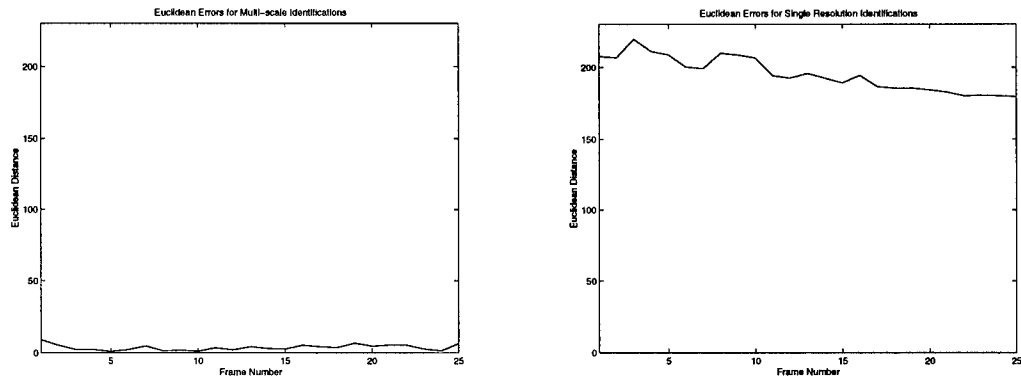


Figure 24. Localization errors for the corrupted jet sequence. Errors were calculated for both the single resolution and multi-resolution identification procedures by computing the Euclidean distance between identified target locations and known position information. The ADP tracking system is capable of identifying the target in noisy imagery and only introduces small errors into the localization measurement. Single resolution identification techniques are not as robust, and the algorithm was not capable of correctly classifying the target in a single frame of the sequence.

### Semi Sequence

The second image sequence used for measuring the performance properties of the anisotropic diffusion tracking system consisted of 74 infrared images of a semi truck in motion. All images had a resolution of 320x240 pixels, and each pixel represented 256 intensity levels. The ADPs were constructed with a 1 of 2 uniform sampling scheme, a gradient threshold ( $k$ ) of 15, and a  $\Delta t$  of  $\frac{1}{4}$ . Figure 25 displays the pyramid constructed for the first frame of the sequence. For the entire sequence, the most significant element of the semi truck was the trailer. Therefore, the root level of the sequence was defined to be the second resolution representation above the original image

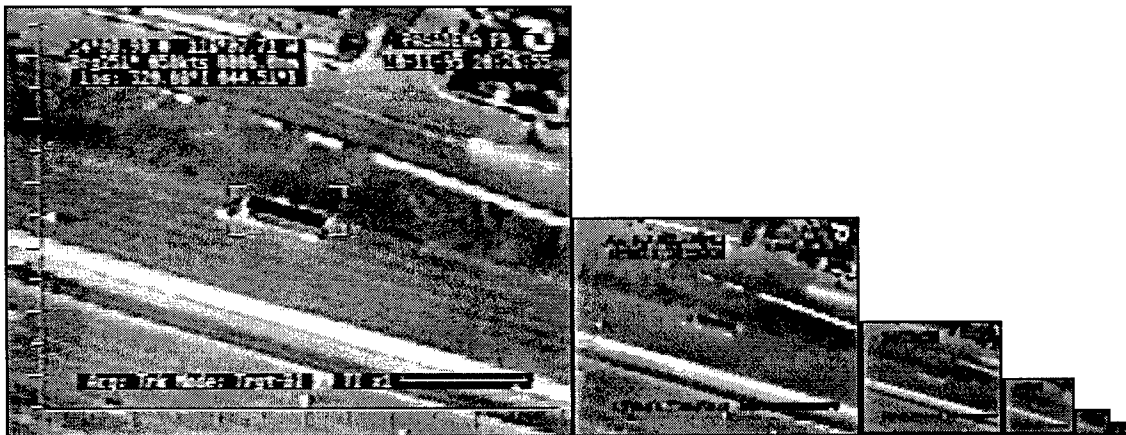


Figure 25. The first frame of the semi sequence and its corresponding anisotropic diffusion pyramid. The semi is visible only in the first three scene representations within the diffusion pyramid, resulting in the selection of the second level of the pyramid as the root level of the multi-scale search.

The first performance measurement of the simulation was the comparison of computational requirements between the multi-scale search and the single resolution identification procedure. Applying the multi-resolution technique to the tracking, the target recognition tasks were performed using a binary template matching routine. For the single resolution match, the algorithm required approximately 46 seconds per frame on a Sun Ultra 1/170, while the multi-resolution technique needed approximately 6 seconds per frame. These results show an overall system performance improvement of 7.7 times tradition single-resolution methods.

While the semi sequence again shows the presence of computational enhancements through the use of the ADP, the performance gains within this sequence account for only 1/3 of those attained with the previous jet aircraft simulation. As these two tracking sequences utilize different pyramid levels for their root level, the differences between the performance improvements within these images displays the sensitivity of the multi-scale method to root level selection. Coarser root levels allow more efficient and robust solutions then finer root levels.

Using the same binary template matching routine, pixel localization errors were also computed for both single and multi-resolution trials. These results are summarized in Figure 26. Localization errors are expressed as the Euclidean distance between the observed point and ground truth. These results show that the increased computational efficiency of the multi-scale search does not introduce extra localization error. For the entire sequence, the multi-scale and single-scale algorithms produce similar measurements. The mean localization error for the anisotropic diffusion tracking system was 1.08 pixels while the mean localization error for the single resolution technique was 1.11 pixels.

To display the robustness of the anisotropic diffusion identification system, the simulations were performed on the same set of images, but corrupted by Gaussian distributed noise. The mean-square signal to noise ratio of the test images was approximately 15.34. As can be seen from the identification results presented in Figure 27, the pixel localization error of the multi-scale technique increased in the presence of noise. However, the algorithm was still capable of estimating the object location in the majority of the frames. The result of the single resolution method was very much in contrast, unable to locate the object during any frame of the corrupted sequence. The ability to find the target in noisy imagery allowed the multi-scale

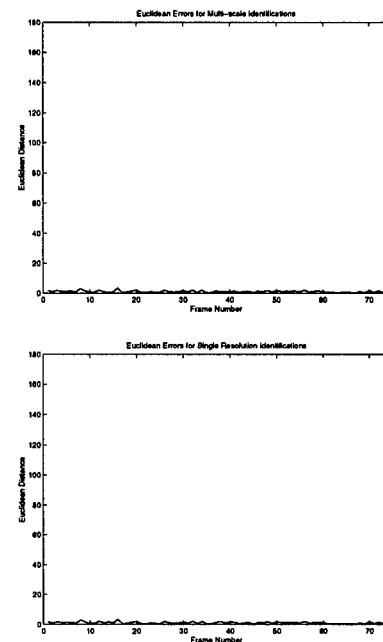


Figure 26. Localization errors for the original semi sequence. Errors were calculated for both the single resolution and multi-resolution identification procedures by computing the Euclidean distance between identified target locations and known position information. The two algorithm produce comparable results, though the multi-scale technique requires less computational resources.

object recognition system to provide a smaller pixel localization error, with a mean error of 19.15 pixels compared to 140.32 pixels of the single resolution system.

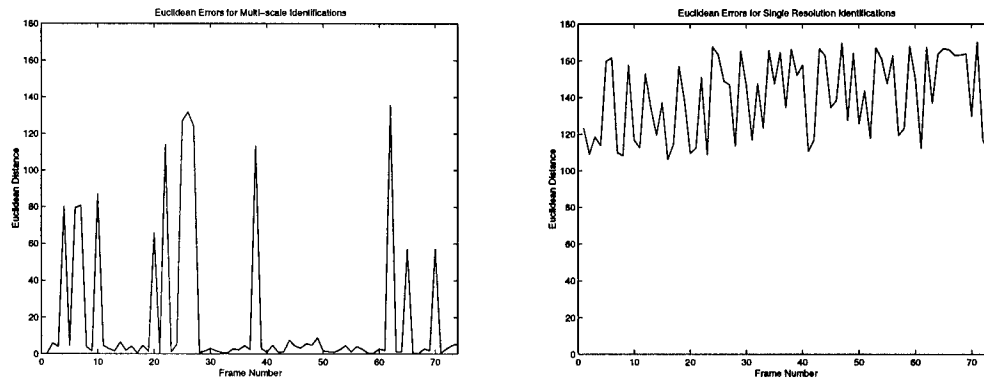


Figure 27. Localization errors for the noisy semi sequence. Errors were calculated for both the single resolution and multi-resolution identification procedures by computing the Euclidean distance between identified target locations and known position information. The multi-scale algorithm is capable of detecting the target in a majority of the frames, denoted by the regions of low measurement error. Single resolution techniques are unable to find the target in any frame, as may be observed by the large measurement error within each frame of the sequence.

### Truck Sequence

The final sequence used in the solution quality simulations consisted of 123 images of the rear of a truck. The original images had a resolution of 320x240 pixels, and each pixel has a range of 256 intensity levels. The pyramids used for this evaluation were constructed with a 1 of 2 uniform sampling scheme, a gradient threshold ( $k$ ) of 15, and a  $\Delta t$  of  $\frac{1}{4}$ . Figure 28 shows the pyramid constructed for the first frame. To implement a multi-resolution search, the root level of the object must be identified. For the entire sequence, the largest element of

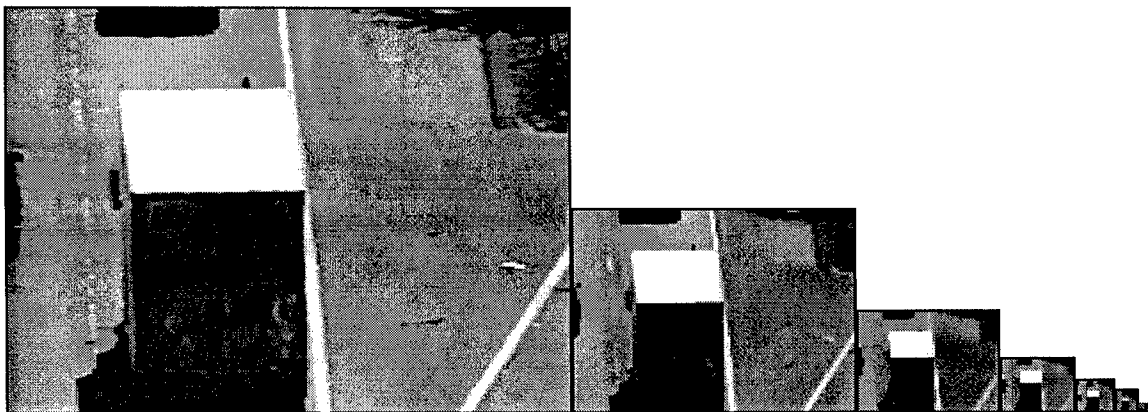


Figure 28. The first frame of the truck sequence and its corresponding anisotropic diffusion pyramid. The truck is still visible in the sixth scene representation within the diffusion pyramid.

the truck is the back of its trailer. Thus, the root level of the sequence was defined to be the fifth resolution representation above the original image. However, presence of other large objects in the scene necessitate selecting a lower initial level for the multi-scale search (in this example, the road and frame edge must be considered objects). In the following simulations, the root level of the target was defined to be the third resolution representation above the original image. Figure 30 shows the multi-scale edge template.

Applying multi-resolution techniques to the object recognition problem and using the third level of the pyramid as the root level, object recognition tasks were performed using a binary, edge based, template matching routine. For a single resolution match, the algorithm required approximately 325 seconds per frame on a Sun Ultra 1/170, while the multi-resolution technique required approximately 9 seconds per frame, including pyramid construction costs. The results show a system performance improvement of 36 times traditional single-resolution methods, again displaying the dependence of the anisotropic diffusion pyramid to the selection of the root level.

Increased computational efficiency does not introduce additional error into the identification results, and using the binary template matching routine, pixel localization errors were computed for both single and multi-resolution trials. These results are summarized in Figure 31, where the localization error is expressed as the Euclidean distance between the identified target location and ground truth. For the first 55 frames, the algorithms produce similar measurements. During the remaining images of the sequence, portions of the truck become occluded, with the top of the truck moving out of the image during frames 56 to 78 and the side of the truck occluded during the rest of the sequence. Both identification techniques are incapable of locating the target when the top of the truck is absent from the frame; however, the

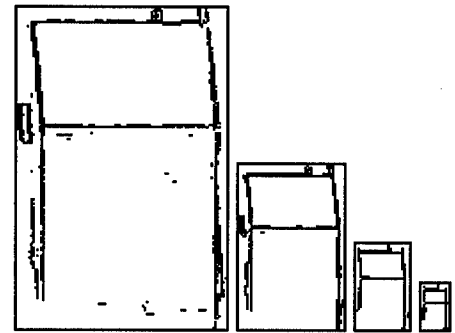


Figure 29. The multi-scale template used for the truck sequence.

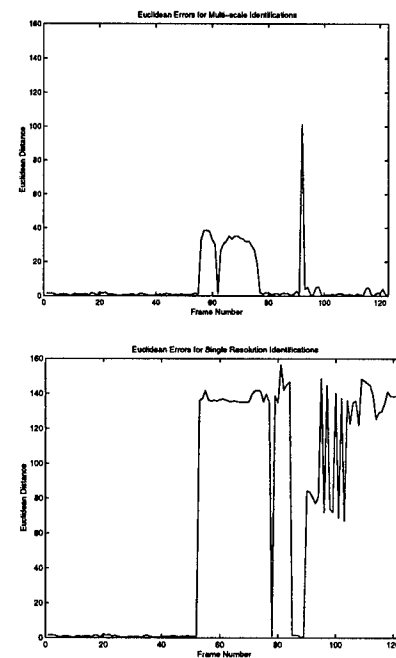


Figure 30. Localization errors for the original truck sequence. Errors were calculated for both the single resolution and multi-resolution identification procedures by computing the Euclidean distance between identified target locations and known position information. The two algorithm initially produce comparable results. At approximately frame 55, significant portions of the truck become occluded. Upon reappearance, the multi-scale approach is capable of identifying the slightly deformed target while the single resolution technique is not.

single-resolution method is also unable to accommodate the occlusion of the side panel in the later portions of the sequence. Anisotropic diffusion pyramids, and their coarse to fine search, are more resilient to these target changes, reacquiring the truck as it becomes entirely visible in the scene. Overall, the multi-scale technique had an average error of 7.06 pixels and the single resolution technique had an average error of 68.31 pixels.

Performing the simulations on a corrupted representation of the image sequence again displays the increased robustness of the anisotropic diffusion pyramid. (The image sequence was created by adding Gaussian noise to the original images, resulting in a mean-square signal to noise ratio of 2.64, and the first frame of the noisy sequence is shown in Figure 31.) As can be seen from the data presented in Figure 33, the pixel localization error of the multi-scale technique increases in the presence of noise, while the single resolution method actually provides better results than attained on the original image set. The ability of the anisotropic diffusion pyramid to provide similar solutions to the identification problem in the presence of noise makes the multi-scale structure a more robust solution to the object identification problem and allows its mean error to increase by only 14.60 pixels. The mean error for the single resolution identification method decreased by 41.60 pixels, providing little correspondence to the original image sequence results.

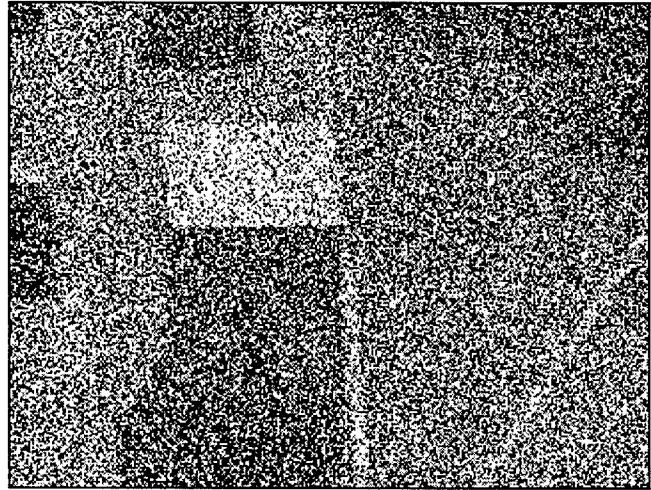


Figure 31. The first frame of the noisy truck sequence. The images were corrupted with Gaussian additive noise.

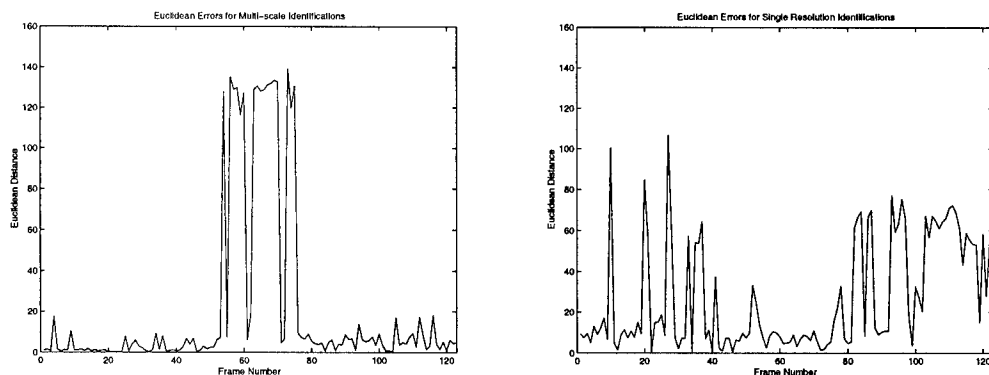


Figure 32. Localization errors for the corrupted truck sequence. Errors were calculated for both the single resolution and multi-resolution identification procedures by computing the Euclidean distance between identified target locations and known position information. Using the anisotropic diffusion pyramid produces similar results between the original and noisy sequences. Application of traditional, single resolution techniques produce significant deviations in identification performance.

### C. PUBLICATIONS

#### Appearing in print and in press:

- [1] S.T. Acton, "Multigrid anisotropic diffusion," *IEEE Transactions on Image Processing*, vol. 7, pp. 280-291, 1998.
- [2] S.T. Acton and A.C. Bovik, "Nonlinear image estimation using piecewise and local image models," *IEEE Trans. on Image Processing*, vol. 7, pp. 979-991, 1998.
- [3] S.T. Acton, "On fuzzy nonlinear regression for image enhancement," *Journal of Mathematical Imaging and Vision*, vol. 8, pp. 239-253, 1998.
- [4] S.T. Acton, "Image restoration using generalized deterministic annealing," *Digital Signal Processing: A Review Journal*, vol. 7, pp. 94-104, 1997.
- [5] S.T. Acton and A.C. Bovik, "Piecewise and local image models for regularized image restoration using cross validation," *IEEE Transactions on Image Processing*, in press, to appear early 1999.
- [6] C.A. Segall and S.T. Acton, "Morphological anisotropic diffusion," *Proc. of the IEEE Int. Conference on Image Processing*, Santa Barbara, Ca., October 26-29, 1997.
- [7] W. Chen and S.T. Acton, "Morphological pyramids for multiscale edge detection," *IEEE Southwest Symposium on Image Analysis and Interpretation*, Tucson, April 6-7, 1998.
- [8] K. Pope and S.T. Acton, "Modified mean curvature motion for multispectral anisotropic diffusion," *IEEE Southwest Symposium on Image Analysis and Interpretation*, Tucson, April 6-7, 1998.
- [9] S.T. Acton, "Anisotropic diffusion and local monotonicity," *IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP-98)*, Seattle, May 1998.
- [10] C.A. Segall and S.T. Acton, "Multi-scale video tracking," *Workshop on the Detection and Classification of Difficult Targets*, Redstone Arsenal, Huntsville, AL, June, 1998.
- [11] S.T. Acton, "A PDE technique for generating locally monotonic images," *IEEE Int. Conference on Image Processing*, Chicago, October 4-7, 1998.
- [12] C.A. Segall and S.T. Acton, "Gradient independent translation via differential morphology," *IEEE Int. Conference on Image Processing*, Chicago, October 4-7, 1998.
- [13] S.T. Acton, "Edge enhancement of infrared imagery by way of the anisotropic diffusion pyramid," *Proc. of the IEEE Int. Conference on Image Processing*, Lausanne, Switzerland, September 16 - 19, 1996.
- [14] S.T. Acton, "A pyramidal edge detector based on anisotropic diffusion," *Proc. of the IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP-96)*, Atlanta, May 7-10, 1996.
- [15] M.C. Burton and S.T. Acton, "Target tracking using the anisotropic diffusion pyramid," *Proc. of the IEEE Southwest Symposium on Image Analysis and Interpretation*, San Antonio, April 8-9, 1996.

- [16] A. Yaicharoen and S.T. Acton, "A hybridized edge preservation coefficient for anisotropic diffusion," *Proc. of the SPIE Symposium on Visual Communications and Image Processing (VCIP-96)*, Orlando, March 17-20, 1996.
- [17] C.A. Segall, W. Chen, and S.T. Acton, "Nonlinear pyramids for object identification," *Proc. of 30th Annual Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, California, Nov. 3-6, 1996.
- [18] C.A. Segall and S.T. Acton, "Sampling conditions for anisotropic diffusion," accepted for publication in the *Proc. of the SPIE Symposium on Visual Communications and Image Processing*, San Jose, January 23-29, 1999.
- [19] C.A. Segall and S.T. Acton, "Adaptive templates in scale-space," accepted for publication in the *Proc. of the SPIE Conference on Real-time Imaging*, San Jose, January 23-29, 1999.

Submitted:

- [20] C.A. Segall and S.T. Acton, "The anisotropic diffusion pyramid for multi-scale tracking," submitted to the *IEEE Transactions on Pattern Analysis and Machine Intelligence*.
- [21] C.A. Segall, W. Chen, and S.T. Acton, "Video tracking using the morphological pyramid," submitted to the *Journal of Electronic Imaging*, under revision.
- [22] S.T. Acton, "Locally monotonic diffusion," submitted to *IEEE Transactions on Signal Processing*.
- [23] S.T. Acton and A.C. Bovik, "Segmentation by locally monotonic reduction," submitted to *Journal of Applied Signal Processing*.
- [24] A. Restrepo (Palacios) and S.T. Acton, "2-D binary locally monotonic regression," submitted to the *IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP-99)*, Phoenix, March 14-19, 1999.
- [25] S.T. Acton and A. Restrepo (Palacios), "Definitions and algorithms for multidimensional local monotonicity," submitted to the *IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP-99)*, Phoenix, March 14-19, 1999.
- [26] J. Bosworth and S.T. Acton, "Morphological local monotonicity," submitted to the *IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP-99)*, Phoenix, March 14-19, 1999.

#### ***D. PARTICIPATING PERSONNEL AND DEGREES AWARDED***

<u>Name</u>	<u>Degree Awarded</u>
1. Andrew Segall	M.S., Fall 1997, Ph.D. expected Spring 2001
2. Timothy Jackson	M.S., Spring 1996
3. Wei Chen	Ph.D., Summer 1997
4. Anthony Wright	M.S., Fall 1998
5. Takashi Koshimizu	M.S., Summer 1998
6. Matthew Burton	M.S., expected Fall 1998
7. David Cary	M.S., expected Spring 1999
8. Zhongxiu Hu	M.S., expected Spring 1999
9. Joseph Bosworth	Ph.D., expected Spring 2001
10. Mark Kargel	Ph.D., expected Spring 2001

#### ***E. REPORT OF INVENTIONS***

No patents or reportable inventions.

## BIBLIOGRAPHY

- J. Babaud, A.P. Witkin, M. Baudin and R. Duda, "Uniqueness of the Gaussian kernel for scale-space filtering," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.8, no.1, pp.26-33, 1986.
- P.J. Burt, "Fast filter transforms for image processing," in *Computer Graphics and Image Processing*, vol.16, no.1, pp.20-51, 1981.
- P.J. Burt, T. Hong and A. Rosenfeld, "Segmentation and estimation of region properties through cooperative hierarchical computation," in *IEEE Transactions on Systems, Man, and Cybernetics*, vol.11, no.12, pp.802-9, 1981.
- F. Catté, P.-L. Lions, J.-M. Morel and T. Coll, "Image selective smoothing and edge detection by nonlinear diffusion," in *SIAM Journal of Numerical Analysis*, vol.29, no.1, pp.182-93, 1992.
- D.A.F Florêncio and R.W. Schafer, "Homotopy and critical morphological sampling," in *Proceedings of the SPIE's 1994 International Symposium on Visual Communications and Image Processing*, Chicago, Illinois, September 1994.
- R.M. Haralick, S.R. Sternberg and X. Zhuang, "Image analysis using mathematical morphology," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.9, no.4, pp.532-50, 1987.
- J.J. Koenderink, "The structure of images", in *Biological Cybernetics*, vol.50, pp.363-370, 1984.
- T. Lindeberg, "Scale-space for discrete signals" in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.12, no.3, pp. 234-54, 1990.
- P. Maragos, "Slope transforms: theory and application to nonlinear signal processing," in *IEEE Transactions on Signal Processing*, vol.43, no.4, pp.864-77, 1995.
- D. Marr and E. Hildreth, "Theory of edge detection," in *Proceedings of the Royal Society of London*, vol.B207, pp.187-217, 1980.
- Miller, J.L. 1993. *Principles of Infrared Technology*. Reinhold, New York.
- Molley, P. 1989. Implementing the difference squared error algorithm using acousto-optic processor. *Proc. Of the SPIE*, vol. 1098, pp.232-239
- Nasr, H. and F. Sadjadi. 1989. Automatic target recognition algorithm performance evaluation: the bottleneck of the development life cycle. *Proc. Of the SPIE*, vol. 1098, pp.156-160.
- W. Niessen, B.M. ter Haar Romeny and M. Viergever, "Numerical Analysis of Geometry-Driven Diffusion Equations", *Geometry-Driven Diffusion in Computer Vision*. Dordrecht: Kluwer Academic Publishers, 1994, ch.15, pp.393-411.

- P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-12, no.7, pp.629-639, 1990.
- W.K. Pratt, *Digital Image Processing*. New York: Wiley, 1978, pp. 495-501.
- A. Toet, "A morphological image decomposition," in *Pattern Recognition Letters*, vol.9, pp.255-61, 1989.
- F. Torkamani-Azar and K.E. Tait, "Image recovery using the anisotropic diffusion equation," in *IEEE Transactions on Image Processing*, vol.5, no.11, pp. 1573-78, 1996.
- D.V. Widder, *The Heat Equation*. New York: Academic Press, 1975.
- A.P. Witkin, "Scale-space filtering," in *Proceedings of the International Joint Conference on Artificial Intelligence*, pp.1019-22, 1983.
- R.Y. Wong and E.L. Hall, "Sequential Hierarchical Scene Matching," in *IEEE Transactions on Computers*, vol.27, no.4, pp.359-66, 1978.
- L. Wu and X. Xie, "Scaling theorems for zero-crossings" in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.12, no.1, pp. 46-54, 1990.
- Y.-L. You, W. Xu, A. Tannenbaum and M. Kaveh, "Behavioral analysis of anisotropic diffusion in image processing," in *IEEE Transactions on Image Processing*, vol.5, no.11, pp. 1539-53, 1996.